马克思主义哲学是统计科学应用的理论基础

戴世光 著

MAKESIZHUYIZHEXUE
SHITONGJIKEXUEYING
YONGDELILUNJICHU

中国人民大學出版社

马克思主义哲学
是统计科学应用的理论基础

戴世光 著

中国人民大学出版社

602.81/37

马克思主义哲学是统计科学应用的理论基础

戴世光 著

中国人民大学出版社 出 版
（北京西郊海淀路39号）
中国人民大学出版社印刷厂印刷
（北京鼓楼西大石桥胡同61号）
新华书店北京发行所发行

开本：850×1168毫米32开 印张：2.75
1985年4月第1版 1985年4月第1次印刷
字数：64,000 册数：1—11,000
统一书号：3011·183 定价：0.47元

前　言

一百年来，在统计科学的发展中，“统计数字”与“统计学”一直属于两个完全不同的基本统计概念。统计数字（现代科学界通常称之为“资料”）乃是科学研究工作（或政府统计工作）通过观测或调查，搜集一定空间、时间下反映客观自然、社会现象的数字资料。统计学即以概率论为基础的数理统计学，乃是应用数学的一个分支；统计学的应用，则为分析客观大量现象数量特征、数量关系、数量变化的一套统计方法，属于科学研究的工具。是以，统计数字与统计学的关系，就是数字与数学方法的关系。统计方法是应用来分析反映客观大量现象的数字资料的工具。一方面是客观存在的大量现象，一方面为分析大量现象的一套数学方法，因此，统计数字与统计学是具有应用关系的，但二者却是不同的两回事。

作者长期从事统计学教学研究实践，在学习马克思主义哲学以来，才逐步认识到：在上列两方面的相互关系中，存在着一个根本性的统计学应用问题，即：科学研究工作分析客观大量现象的数量方面，何以需要一套以概率论为基础的应用数学方法？而客观的大量现象又与概率论的数学原理之间具有何种科学的应用关系？作者认为：这是统计科学在其应用方面，必须解决的哲学理论问题。

从统计科学发展的历史来看，统计学在各个科学领域中的广泛应用，经过长期实践，取得了大量的重要而显著的成就。近代以来，大量事实充分证明：现代统计科学的确是科学研究、科学

实验、科学管理的一种必要的有效工具。但是，统计理论本身和统计方法应用的大量成果，都不能从哲学理论的高度，对前面提出的统计学应用的根本问题，作出科学的理论解释。应该说，在现代统计科学的发展中是缺乏哲学理论作为统计学应用于分析客观存在的大量现象数量方面的指导思想的。

这本《马克思主义哲学是统计科学应用的理论基础》专著，是作者在学习马克思主义哲学的基础上对现代统计科学应用的根本问题的一项统计理论研究。由于作者科学水平有限，缺点、错误难免，敬祈读者批评指正。

戴 世 光

1984年6月于

中国人民大学

目　　录

前　言……………………………………………………………… I
一、应用马克思主义哲学揭示统计科学发展中的若干理论问题……………………………………………………………… 1
二、统计学的基本问题：什么是统计科学研究的特殊对象……………………………………………………………… 9
三、统计学的基本方法：大量观察与平均法……………………20
四、统计学的基本理论：大数定律与概率分布…………………33
五、辩证唯物主义是统计科学应用的理论基础…………………49
六、"实事求是"是应用统计方法的指导思想…………………63
七、具体地分析统计方法在科学研究客观规律性中的具体作用……………………………………………………………71

一、应用马克思主义哲学揭示统计科学发展中的若干理论问题

1. 马克思主义哲学与统计学

作为马克思主义哲学的辩证唯物主义（或唯物主义辩证法）是建立在全部现代科学基础上的唯一科学的哲学理论。马克思主义哲学的研究对象系世界上一切运动和发展的最一般规律。而各门科学的研究对象则为客观某一领域或某方面的特殊规律。马克思主义哲学是要揭示作为一切现象和过程的基础的共同性的东西，提供关于支配任何运动和发展的一般规律的知识，不管运动和发展是发生在自然界中、社会中、还是人的思维中。同时，马克思主义哲学不仅研究整个世界中发生作用的最一般的规律，而且研究这些一般规律在认识过程中的运用。所以，作为马克思主义哲学的辩证唯物主义既是科学的哲学理论，同时又是观察、分析、研究及处理各种科学问题的方法论。

毛泽东同志指出："事物的矛盾法则，即对立统一的法则，是唯物辩证法的最根本的法则。"①是以，辩证唯物主义的基本规律就是对立统一规律，它揭示出事物发展最根本的问题：源泉和动力是什么的问题。各个科学领域，如物理、化学、生物、社会的发展变化，同样地受这一基本规律的支配，种种运动都是由于内部矛盾发展的结果。世界上一切事物的内部都有矛盾性，从而就引起世界一切事物的运动和发展。对立统一规律之所以是辩证唯物主义的核心，就因为它是唯物辩证法的规律和基本范畴的

① 《矛盾论》，《毛泽东选集》第1卷，第274页。

基础。规律与范畴都是反映事物本质的，而规律则是更深刻地反映事物之间必然的本质的内在联系。这里，除了对立统一规律以外，可以列举的唯物辩证法的规律，有：量变与质变的规律，否定的否定的规律等。唯物辩证法的基本范畴则有：现象与本质，必然性与偶然性，原因与结果，质与量，形式与内容，有限与无限，肯定与否定，可能性与现实性，必然与自由等。应该说：对立统一规律是唯物辩证法中其它规律和各个基本范畴的基础和内容，而唯物辩证法的其它规律和各个基本范畴则是对立统一规律的表现和展开。

哲学的范畴与各门科学的范畴不同，哲学的范畴是适用于一切科学的最一般的概念。任何一个科学家，不论他是物理学家、化学家、生物学家、社会学家、经济学家，或者其它专家，都不能不使用下列一些最一般的概念和范畴，如：规律、规律性，矛盾，本质与现象，必然性与偶然性，原因与结果，质与量，形式与内容等等。这些范畴反映出现实世界中各种现象之间的最一般的内在联系；而且它们也是认识世界的阶梯，思维的工具。哲学的规律和范畴概括了人类研究世界的历史经验。统计科学属于方法论性质的科学，乃是科学研究的一套方法、技术（或工具）。作为科学研究的工具的统计科学，首先必须联系“现象与本质”这一基本范畴，因为科学的任务就是透过现象认识本质的。同时，应用统计学必须联系“必然性与偶然性”这一基本范畴，因为，作为统计学本身的理论基础的概率论，就是研究偶然性（也即随机现象）的抽象的数量规律性的。

作为马克思主义哲学的辩证唯物主义，或唯物辩证法，乃是科学的世界观和方法论。由于辩证唯物主义完全是从一切科学的研究成果中概括出来的，因而，唯物辩证法又是一切科学的方法论，成为指导一切科学研究和实践的理论基础。应该说：哲学和其它各门科学都是以客观世界为研究对象的，都要从对客观世界

的研究中，达到关于客观规律性的认识。哲学研究客观世界总体的一般规律性，各门科学研究客观世界某一领域、某一方面的特殊规律性，如数学的、物理学的、化学的、生物学的、社会科学的等等具体规律。哲学不能为各门具体科学提供现成的结论，但是，一切科学在其科学研究实践的过程中，不论有意识地或者无意识地进行科学探索，都离不开辩证唯物主义哲学理论的指导。当前，在哲学与科学的关系上，一方面要批判那种以所谓“科学的科学”的“哲学”代替具体科学作结论，鼓吹“马克思主义哲学就是自然科学基础理论”的“代替论”，同时也要反对那些撇开世界观、方法论的指导，主张“科学无须求助于哲学”的实证论。①

统计科学即数理统计学，为应用数学的一门分支，是科学研究的一套方法、技术。统计科学应用来作为各门科学研究的一种必要工具，也离不开作为马克思主义哲学的辩证唯物主义的指导。所以，辩证唯物主义乃是应用统计学（即数理统计方法，或统计方法）的理论基础。它对应用统计方法于经济学、社会学、教育学、心理学、生物学等等领域的科学研究在探索客观规律性的实践活动中，具有哲学普遍原理的指导意义。

2. 现代统计科学发展中的若干统计理论问题

十九世纪中叶，比利时的数学家、天文学家、物理学家、人口学家、人类学家及社会统计学家，后来科学界称为“统计学之父”的凯特勒（1794—1874），先研究天文、气象方面的统计资

① 见周培源等《以马克思主义哲学作为指导加速发展我国的科学技术》，《光明日报》1978年8月25日。

料，后又利用统计数字研究植物界和人类社会。他在四十年间写了大量的统计著作，主要的有：《比利时王国的统计研究》、《人间论》、《统计研究》、《概率论书简》、《社会物理学》（包括“平均人”理论）、《人体测量学》等。从而他通过统计科学研究的大量实践，在概率论的基础上，把几派统计学术活动中的统计方法，融合成为一门方法论性质的统计学。一百余年来，作为科学研究方法(或工具)的统计科学（即数理统计学）经过理论研究和科学实践，有了很大的发展。然而，若根据马克思主义哲学理论，对统计科学的理论和实践进行辩证唯物主义的分析，就不难看出：在现代统计科学发展的历程中，还存在着若干有待解决的统计理论问题：如方法论性质的统计科学的研究对象问题，理论统计学与应用统计学之间相互关系的问题，理论统计学中的概率论与数理统计这两部分之间的关系问题。除去这些问题之外，还要特别指出：作为现代科学研究方法的统计科学，虽然在科学实践中得到了广泛的应用，但从应用的指导思想来看，还存在着从来没有被注意到的科学与科学之间的理论关系问题。这个问题包括两个要点：一是数学的概率论与客观现象数量情况之间的理论关系问题；二是作为一种科学研究方法(工具)与认识客观规律性的唯物辩证法之间的理论关系问题。现在，对于上列几个统计理论问题的解决，就要进一步问：究竟是仅仅依靠在科学实践中的实际证明呢？还是同时需要在实践过程中根据马克思主义哲学，即辩证唯物主义哲学的理论指导呢？作者认为，这是应该从哲学理论的高度，联系统计科学和实质性科学（如物理、化学、生物学、社会学、经济学等）的研究实践，进行科学研究的一个重大课题。

据此，下面将具体阐明有待研究的五个主要统计理论问题。

（一）统计学的研究对象问题。

现代统计科学，即数理统计学，乃是应用数学的一个分支，

属于一门方法论性质的科学。一百余年以来，作为科学研究方法的统计学，大量应用于科学研究的实践，证明了它是科学研究、科学实验、科学管理的一种必要的方法技术、一套对统计数字进行统计分析的工具。然而，我们知道：科学是客观特殊规律的系统反映，理论是关于客观特殊规律的具体认识。统计学既是一门独立的科学，它就必然具有特殊的研究对象及其特殊的规律性。这就要进一步提出问题，统计科学作为应用数学一个分支所研究的特殊对象，究竟属于什么样的抽象的数量关系及其数量规律性呢？同时，当把统计学这一套方法技术应用于客观实质性的科学研究时，究竟又具有什么样的特殊对象呢？在研究实质性的具体规律性方面又起到什么样的特殊作用呢？这在统计理论研究中就成为一个到现在还未曾提出、更未曾解决的统计科学研究的对象问题。

（二）理论统计学与应用统计学的关系问题。

现代统计科学既是应用数学的一个重要分支，它本身必然是与其特殊的研究对象和数量规律性有关的一套数学原理（如概率论、相关理论等）。但是，作为应用数学而论，这一套关于统计学的数学原理又是用于科学研究对统计数字进行统计分析的一套方法、技术（如平均法、平均数的估计和检验等）。我们知道：马克思主义哲学从其理论而言，乃是辩证唯物主义；就其应用来说，则为唯物主义的辩证方法。数学科学也是如此，从其理论说乃是数学原理；就其应用言，则为数学方法。所以说：数学是一切科学的工具。正是由于这一点，现代统计科学才在分科方面首先分为理论统计学（或统计理论）和应用统计学（或统计方法）；进而在应用统计学方面又按其应用的科学领域，分为社会统计学、经济统计学、教育统计学、生物统计学、医疗统计学等等。于是，应用统计学的统计研究工作，很自然地按照统计方法应用于不同的实质性科学，如经济学、生物学、物理学等科学领

域来区分。但是，在上述统计科学研究实践的条件下，就产生了一个理论与实践脱节的问题。数学领域的统计科学研究工作者就埋头专门研究统计理论方面的数学理论问题，抽象地创造并发展理论统计学；相反的，在应用统计方法于经济、生物等科学领域的统计研究工作者，却面对经济、生物等研究对象只管“用方法”，“把数字代入统计公式”取得数据，并对之直接作出经济学或生物学的解释。这样，从一门独立的统计科学的发展而论，实际上就存在着：未从哲学理论高度来解决理论统计学与应用统计学二者相互之间的理论关系问题。

（三）概率论与数理统计二者之间的关系问题。

作为现代统计科学的理论方面是理论统计学，数学性质明确，内容则为应用数学分支的一套数学方法原理。但是，许多统计学家对于理论统计学的两个基本部分：概率论和数理统计二者的相互关系，往往出现理解混淆不清的问题。例如，美国赫格所著《数理统计学导论》①一书的基本部分主要包括：概率分布、估计和检验等三部分，而其中概率分布理论约占一半以上。然而，我国复旦大学编的《概率论》（1979年）一书则分为：概率论基础、数理统计和随机过程三部分，而对于数理统计则又列成：抽样分布，假设检验和线性统计推断；最小二乘法，四归分析和方差分析。以上列两本统计理论著作为例，在基本内容的体系方面显然是混淆不清的。据此理应指出：现代统计理论界，不仅尚未从哲学理论高度认识统计理论的基本体系，不明确理论统计学的两个基本组成部分：概率论和数理统计的基本内容；而且更未认识到：概率论与数理统计二者相互之间的理论关系。

（四）数学的概率论与客观现象数量情况之间的理论关系问

---

① Hogg and Craig, «Introduction to Mathematical Statistics», 1970.

题。

现代统计科学的应用方面为应用统计学，乃是理论统计学的数学方法原理在客观各个科学领域的科学研究工作中的具体应用。举例来说：在实际的经济统计研究工作中，将常态分布应用来配合大量商品价格在两年间的变动率的次数分配；或者应用算术平均法求得某一工厂全部生产工人月工资次数分配的平均工资水平。这里，配合常态分布曲线则系概率论分布在价格统计研究中的具体应用；运用算术平均法实质上乃是数理统计方法在工资统计研究中的具体应用。表面上这些不过是应用统计方法来分析客观经济现象的数量规律性的计算例，但在应用中实际却存在着：以概率论为基础的统计方法与客观现象数量情况之间究竟具有什么样理论关系的问题。如上例，常态分布在性质上属于一种抽象的随机变量的分布，现在应用来描述价格变动率的次数分配。这究竟是二者在形式上相同呢？还是抽象的概率分布与客观存在的价格变动率分配之间有一定的理论联系呢？再如，算术平均的数学方法具有使次数分配的数量差异相互抵消的作用，现在应用算术平均法来平均全部生产工人的工资差别。这里，抽象的数量差异的相互抵消与客观存在的工人间工资差异的相互抵消，这两种差异之间究竟具有什么共同一致的理论意义呢？应当着重指出：通过上列统计学具体运用实例所提出的问题，实质上是抽象的数学概率论的具体应用与客观实际现象的数量情况之间，究竟具有何种科学理论关系的问题。这一统计理论上的基本问题，如果不从哲学理论的高度得到科学的解释，就无从理解应用统计学于分析研究客观存在的数量规律性所根据的科学理论关系，也难以明确应用统计学这一套工具在工业产品质量控制、农业田间试验、经济统计预测、经济管理决策等等科学研究实践中，所起的具体分析的作用。

（五）统计方法与探索客观规律性的唯物辩证法之间的理论

关系问题。

国际统计学界往往从科学性质上指出统计学（即数理统计学）乃是科学研究的一种方法（工具）。这一点是应该肯定的。因为科学研究的任务就是为了探索客观存在的规律性，只有在科学研究工作中运用科学的方法，才能获得有关某一客观具体规律性的科学结论。然而，现在的问题是：作为科学研究方法的统计科学在具体应用中，究竟可以得到关于客观具体规律性的什么样的数量结论呢？这种具体的数量结论又与要探索的客观具体规律性之间，有什么样的理论关系呢？对于这一科学基本理论问题，现代理论统计学家和应用统计学家都从未考虑过，更谈不上作出任何有关的答案。应当指出：如果不从哲学理论的高度来理解，就无从认识到作为科学研究方法（工具）的统计学，在探索客观规律性中的具体作用与其它科学研究方法（如唯物辩证法，统计调查方法等）相互配合的理论关系。

综合以上有关现代统计科学发展中的五个统计理论问题，可以看出，显然是不能只依靠统计科学本身的数学理论解决的。应该说：上列问题，在理论上属于科学与科学之间的关系问题，在科学实践中则涉及人们的主观认识与客观存在之间的关系问题。如此重要的科学理论问题，只有提高到马克思主义哲学理论的高度，实事求是地进行科学研究，才能取得比较正确的科学结论。问题就是矛盾。只有以马克思主义哲学理论作为我们研究统计科学实际应用的指导思想，具体地分析和研究统计科学理论联系实践中的主要矛盾，统计科学才能在更坚实的理论基础上，得到有效的发展。

二、统计学的基本问题：什么是统计科学研究的特殊对象

从辩证唯物主义的基本观点看来，统计学与统计资料是两回事。统计学是一门方法论性质的科学，统计资料是反映客观社会、自然现象的统计数字资料，统计数字不能视为科学。但是，统计资料又与统计学具有密切的关系，统计数字是应用统计方法（统计学）处理、分析、研究、探索大量现象数量规律性的对象材料。

历史上，早在奴隶社会就产生了为了国家政治、经济管理的需要，由政府记录国家基本情况（人口、土地）的统计资料的工作。到了资本主义社会，由于工业发展对于科学研究的需要，才产生了作为一门科学方法的统计学。统计学是一门应用数学，统计理论与统计方法为统计学本身密切联系的两个方面。现代统计方法（也即以概率论数理统计为基础的统计理论的应用方面）已经成为科学研究、科学试验、科学管理的重要工具（方法、技术）。人们说："统计学是获得科学知识的一套方法"，或者说："统计学是一种科学研究方法"，正是就统计学作为统计方法来概括统计科学的基本性质的。然而，一切科学研究，既然已有数学方法作为科学工具，何以还另外需要一门作为应用数学分支的统计方法呢？从科学理论看，这个问题实质上是统计科学研究客观何种特殊矛盾（或统计学特殊的研究对象，或统计方法应用的特殊对象）的基本问题。科学的发展产生于人们观察客观现

象，研究客观规律性的需要。没有研究客观某种特殊数量规律性的需要，就不会产生一门特殊的作为应用数学分支的统计学。

根据辩证唯物主义的基本观点，要探讨什么是统计学的基本问题，就需要研究这一问题的历史和现状，即统计科学产生和发展的历史和现状。为此，现在先由“统计”发展成统计科学的简史来考察特殊需要一门统计科学的基本原因。“统计”成为一个科学概念，就是首先来源于人们了解国家基本事实的需要。例如，古代政府记录全国人口、土地情况的工作就属于政府统计工作的原始状态，对情况所记录的材料则系编制统计数字的重要依据。在统计历史中，德国的记录学派，是从封建主义后期，由大学教授康令著书名为《现时政治上最重要事件》（1660年）作为讲授“当前欧洲学”的讲义开始的。1749年，又由德国教授阿罕瓦尔进一步扩大“国家学”的范围，用叙述体记录各国国情，称之为“政治家所应知之事”；并且将之写成著作，起名为“统计”。从此，就用记录学派首创的“统计”一词来表示国家基本国情（人口、土地、物产、兵力等）的叙述和比较。其实，记录学派乃是比较国家基本事实资料的学派（统计史中有人称之为旧派）。显然，如同记录、国势、图表这一学派那样，所统计表列的一个国家一定时间下种种基本事实的统计资料，根本不能成为什么科学。十六、十七世纪的英国，是最早从封建制度向资本主义转变的国家。英国葛兰德于1662年写出《关于死亡表的自然和政治的观察》一书，发表人口出生、死亡数字统计研究的结果，阐明大量人口现象的出生性比例和死亡率的数量规律性；从方法上，他还通过大量观察的实践，首倡“大量恒静”的观点。后人称英国葛兰德及配弟（写《政治算术》一书，1690年出版）为政治算术学派的创始人。他们虽未创立“统计学”之名，但所用以探索数量规律性的方法，则具有“统计学”之实（后来有人称之为新派）。从早期的数理统计理论而言，在数学家们长期研究赌

博取胜的数学概率之后，法国数学家白努里于1713年发表著作，专论作为现代统计学基础的概率原理；法国达姆瓦研究概率的数学理论，于1733年推导出统计学中最重要的公式：常态曲线公式（即被称为大量观察下的常态定律）。到了十九世纪中叶，比利时统计学家，也就是后来科学界称之为“统计学之父”的凯特勒，以概率论为基础，采用对集体（或集团）现象进行大量观察和综合平均的方法，来研究若干自然、社会大量现象的数量规律性，从而把记录、政治算术、数理统计三派观察大量现象，进行数量分析的方法，融合成为一门统计方法的科学，也即发展到现代成为科学研究、科学实验、科学管理必要工具的统计科学。

根据以上现代统计科学的历史发展，初步看出，统计学作为一门独立的科学，乃是产生自研究客观集体大量现象数量规律性的需要。现在，再进一步就统计科学发展的历史过程中，从过去到现在，在应用统计学（以经济统计学为主）和理论统计学（即数理统计学）两个方面，选择十多本重要著作，对其有关统计学定义、性质、对象所作的论述，作比较具体的分析。

（一）凯特勒在1836年发表的《社会物理学》中曾写道：“社会现象之中是有规律性存在的。这种规律性不能先验地决定，唯有求助于实验的研究方法”，而“这种实验的研究法不得不诉之于大量的现象”。他还说：“我们基于概率的计算，就可以发现法则。”这里，凯特勒并没有对统计学下定义，只是总结了他对集体大量现象进行统计科学研究的实践经验。然而在客观上，他却通过科学研究实践，指出了统计科学的发展方向。

（二）英国统计学家鲍来在他所著《统计学原理》（1901年）中，在谈到“统计学方法”时，就指出：“遇有大量数字，描述集体的大数，以及总计数或平均数数列时，均必须采用依据大数性质，能描述复杂集体的特殊方法。”他在谈到“统计学与经济学”时，提到“作为大数科学的统计学”；在谈到“统计学与

社会研究”时，还说：“统计学可以确切地称之为平均数的科学”。英国统计学家游尔和肯德尔在1910年写的《统计理论导论》一书中，则直接概括为：“统计学为包括一系列用于解释说明受大量复杂原因影响的数量资料的统计方法”。可以说，鲍来给予统计学最概括的定义，即“大数的科学”和“平均数的科学”；而游尔则指出：统计学是解释说明大量现象的数量资料的统计方法。可见，他们对统计学的论述，都立足于“大量”和“大数”。

（三）奥国统计学家齐泽克在1913年写了一本《统计平均数》的著作，他在〈导言〉中肯定了英国经济学家埃奇沃思（1845—1926）和鲍来关于“统计学为平均数的科学”之后，根据他在二十世纪初对大量观察的认识，突出地肯定平均数在统计学中的重要地位，指出：“统计学的任务为从一般方法的角度研究平均数的应用和利用，并且确定平均数在统计方法中的地位。”

（四）作为促进现代统计学重大发展的英国统计学家费煦，在《研究工作者用统计方法》（1934年第五版）一书中写道：“统计科学基本上为应用数学的一个分支，可以视为应用来分析观察资料的数学方法。”他说：“因而我们要考虑统计学主要内容的三个方面，并且用数学语言说明所有情况下出现的同一类型的统计问题。为此，统计学可以视为：（1）现象全域（总体）的研究，（2）差异（或变异）的研究，（3）简化资料方法的研究。”“统计学所研究的为全域，也即对全部个体的综合研究，而非对个体的研究，因为科学理论包含有集体大量的综合特征，而非个体的特征。”他又说：“现代统计学家与其前辈在统计学研究差异这个问题上的突出区别……后者除去测定集体综合的或平均的数值以外，就没有其它目的”；“人们熟悉常态样本平均数的误差曲线已有百年，而一直到1915年才以标准差作为学术研究的目标……而现代的观点，还要着重研究任何变数现象差异

的原因”。这里，费煦在应用概率论数理统计方法的基础上，不仅强调了统计学对全域集体大量现象个体差异的研究，而且指出了现代统计学业已从“平均数的科学”发展到“标准差（分析差异的工具）的科学”。

（五）美国经济统计学家密尔士在《统计方法》（1955年）一书中说道：“这本书探讨一种研究的方法：一种考察社会和自然过程并为科学研究和行政管理提供决策根据的方法。从细节看，统计方法技术是种类繁多，彼此不同的；但从总的来看，它们都构成了考察人类情况和自然规则的一个完整的、有系统的、合乎逻辑的研究方式。”美国经济统计学家中国人邵品剡在所著《工商业统计学纲要》（1977年）一书中写道：“大量的数学资料引起需要一套能够用以有效地组织、表列、分析和解释这种大量数字的系统方法。正是由于使用这些方法，才可以获得正确的结论并作出合理的决策。统计方法在根本上就是发展来满足这种需要的。”可以看出，以决策方法为重点的现代经济统计学，依然从概率论出发，强调需要用分析、解释客观集体大量现象数字资料的统计方法。

美国经济统计学家梅森在《工商业和经济学中应用的统计方法》（1978年）一书中，对统计学下的定义为：“为了对客观偶然性差异情况，作出更好决策的目的，而从事搜集、组织、表列、分析、解释差异数量的科学。”

以上三位经济统计学家都指出：（1）客观大量现象导致采用一套组织、表列、分析和解释这种集体大量数字的统计方法；（2）现代统计学是依靠概率论研究社会和自然界偶然性差异，提供决策根据的科学方法。

（六）澳大利亚经济统计学家卡梅鲁在《经济学家用应用统计学》（1977年）一书的〈统计学的本来性质〉一节中写道：“分析统计资料的方法是统计方法……。作为统计方法的根据

的，则为统计理论或数理统计学。统计方法的应用非常广泛，已在生物科学和农业科学方面取得了极大的发展。就统计方法的本来性质而论，它特别适用于处理和分析那些属于差异性的资料。人们对于这种差异性根本不可能用实验方法予以完全控制，只能通过观察来研究可能存在的全域中的一个小部分。显然，如果某一现象表现得完全一致，人们对之只需要作单个的观察，就可以用来发现它的本来性质。然而，如果现象表现为不能用实验控制的差异性，人们就必须发明并设计一套方法技术，用以推论全域的本来性质，或者利用已有的专门观察来测验假设。”此外，卡梅鲁还在该书〈经济学与统计学〉一节中，首先说明：“统计方法能够使经济学的经验方面得到发展，因为应用了统计方法，就能对经济理论的形成，提供真正的事实内容。”继而，他在谈到经济统计资料时，就提到：“经济统计专家的主要任务之一是计量的任务，例如计量‘真实收入’的动态就属于一项困难的计量工作，……而经济统计资料则系用以检验纯粹经济理论的假设，并从这些假设来取得结论的。”

据此，卡梅鲁在这里实际上谈到了三个重要统计问题。第一，他指出了客观现象存在着数量方面的差异性，这是观察的客观前题，如果现象表现为完全一致，就只需要作单个的观察；第二，他说明了统计理论（即数理统计学）与统计方法之间的关系；第三，他分析了经济学与统计学（统计方法）之间的关系，并且还说明了计量经济动态（如“真实收入”的动态）的经济统计数字与经济理论检验假设需要应用统计方法分析实际数据的关系。

（七）美国工商业管理统计学家拉平在《现代工商业决策用统计学》（1982年）一书第一章〈导言〉中，首先提出统计学的定义，他将之概括成为：“统计学为一整套的方法和理论，是在面临着不肯定性的情况下，应用来为作出决策提供数量证据。”他

在这里明确地指出：第一，客观存在着偶然性（即不肯定性）；第二，工商业管理必须在面临着偶然性的情况下作出决策；第三，为了作出决策就必须应用以概率论为理论基础的统计学。但是，他却不明确统计理论与统计方法二者之间的相互关系，更未曾想到：概率论与客观不肯定性之间的关系。

继而，拉平在〈导言〉中还提出统计学的两种两分法，一为："描述统计学和推论统计学"；二为："演绎统计学和归纳统计学"。第一种两分法是从统计学的内容来说的，而第二种两分法则系就应用统计方法进行科学研究的推理方式而论的。应该说，拉平提出统计学的这两种两分法，对于运用现代统计方法，还是具有实际意义的。

（八）日本统计学家山根太郎在所著的《统计学》（1978年）一书中写道："在这大量的信息（按：指资料）当中，如果存在着说明社会现象与自然现象的所谓'法则性'或'规律性'的话，对于社会现象来说，就可以根据了解到的规律性为制订政策提供依据，而对于自然现象来说，就可以发现控制它们的手段。"继而，他又探讨："由社会现象或自然现象构成的大量信息中，是否存在法则性或规律性呢？〈概率论〉这一章和〈正态分布〉一章，说明在大量的数据中，可以作为概率现象理出这种法则性或规律性"；"另外，由概率理论可以导出种种统计方法，由这些统计方法也可以整理与掌握到作为统计现象（按：指大量数量现象）的上述法则性与规律性。"他指出："这样一来，统计学在社会现象或自然现象的分析方面就成了极为重要的分析工具。在经济学与管理学中，由于有着货币这种计算单位，所以作为分析工具的统计学是重要的。并且，在应用到这些领域中充分发挥着它的作用。"

山根太郎写的是经济和管理用的《统计学》，他在阐明社会和自然现象通用的统计方法（分析工具）的基础上，强调运用统

计方法处理集体大量差异的资料，分析大量概率现象规律性的重要作用。英国费煦只提到抽象的“数量差异”，而山根太郎却指出了“统计现象”的“法则性与规律性”，应该说，这在科学认识上，业已有所提高。

（九）香港的中国经济统计学家胡孝绳在《统计学》（1976年）一书中写道：“统计学系以搜集、显示、分析及解释等等统计技术为手段，而达到明白统计群体（按：指作为统计学应用对象的各种社会的、自然的或实验的集体大量按质归类的数字事实）表征（按：指特征）的真相”的科学方法。他并且指出：“统计群体的各个体，除具有此共同性质之外，并各有不同的性质；故可在共同性质之下，观察其不同的性质。此等性质是为统计表征。”胡孝绳在这里不仅说明统计学是观察、搜集、表列、分析、解释大量集体的统计数量特征的科学，而且指出了：大量集体现象中存在个体的共性和个性，统计学乃是分析大量现象共性的数量特征的科学方法。应该强调：这对客观大量现象的个性与共性的统计研究，在科学认识上是进一步深入了。

（十）美国统计学家穆德在《统计学导论》（1963年）一书〈统计学〉一节中写道：“统计学能够分为两大类：（1）描述统计学，研究简缩数据和描述这些数据；（2）推论统计学，研究利用数据去作出决策的方法，即根据部分数据去推断更为一般的情况。这一统计推论的过程大都是依赖包含在样本数据中的不完备信息，它必然要冒风险。原因在于不肯定性，而概率就是这种不肯定性的计量。”继而他又说：“统计学的另一划分，即理论与方法的划分。统计理论是应用数学的一个分支，它扎根于纯数学的一个领域：概率论；事实上，统计理论的完整结构从广义来说，可以认为是包括概率论在内的。”穆德在谈到统计学的应用时，还说：“我们将把统计学看成为科学方法的技术工具。当不肯定的状态普遍存在时，统计学能够提供一套作出决策的工具。

这些工具有着十分广泛的应用，在各个科学领域，如物理学、生物学、社会科学等等，都是用得上的。它们不仅能用在科学界，而且还能用在工商业以及日常生活之中。另一方面，还可以为专门的研究领域具体地设计出某些工具。”

穆德写的这本书在性质上属于“统计学的理论”，对于整个统计科学的两种划分：理论统计学和应用统计学；描述统计学和推论统计学，阐述是比较清楚的。他明确地提到：统计理论方面的概率论和统计方法所面临着普遍存在的“不肯定的状态”。这是他能够把数学的概率论与客观存在的“不肯定性”（按：即偶然性）在实证中直觉地联系起来。但是，穆德不懂得从哲学理论的高度来解释概率论与客观“不肯定性”之间的关系，实质上乃是“认为科学无须求助于哲学”这一错误观点的反映。

（十一）苏联数理统计学家温特切勒在所著的《概率论》（1958年）的〈概率论的对象〉一节中写道：“概率论是一门研究随机现象的规律性的数学学科”，继而并对随机现象作出解释：“在各种物理问题和工程问题的科学研究中，常常会遇到一种特殊类型的现象，即所谓随机现象。随机现象就是在重复实验中有时出现有时不出现的现象。”(书中举例如：和地平线成一定角度的大炮发射出炮弹，同一物体放在天秤上称重多次所得的结果彼此具有的差异性，等等）。温特切勒在举了实际例子之后说：“实践证明，观察了大量的同类随机现象后，通常总揭露出一种完全确定的规律性，也就是大量随机现象所特具的一种稳定性。”最后，他总结：“概率论的数学定律，是客观存在于宇宙中的大量随机现象的统计规律性的反映。概率论应用数学方法来研究这种现象，且由于它的特有方法而使它成为一个数学分支。”

温特切勒对概率论的论述是非常明确的。他不仅明确地提出来“概率论的对象”是从数学学科方面“研究随机现象的规律性”，而且还从实践方面指出：大量观察的随机现象，能揭露出

一种稳定的规律性。这是苏联数理统计学家的科学认识水平比欧美等国家数理统计学家较高的具体体现。

但是，温特切勒同时在〈概率论的对象〉一节中还另外写道：“科学上的概率论方法或统计方法并不是与严正科学的、古典的、常用的方法对立，而是它的补充，使之能根据随机现象所固有的随机因素而对现象作更深入的分析”；并说：“在一切知识领域内，广泛而有效地应用统计方法，是自然科学和技术科学发展的现阶段的特征”。据此，又看到苏联数理统计学家温特切勒对于现代统计科学的理解不足之处。他不仅含混其辞地把概率论与统计方法等同起来，未曾区别概率论与数理统计；而且又把统计方法的应用局限于自然科学和技术科学。这种情况也可以说是：由于未能从哲学理论的高度来认识现代统计科学与自然、社会等实质性科学之间的理论关系，从而错误地把自然科学与社会科学对立起来。

（十二）我国数理统计学家林少宫在《基础概率与数理统计》（1963年）一书中写道：“概率论与数理统计是一门研究偶然现象的规律性的科学”；“这种现象是属于偶然性质的，我们不能用一些简单的物理定律加以概括，而须从大量观测中综合分析，归纳出一些‘大量现象’的规律来。”近来，我国复旦大学数学系编写的《概率论与数理统计》（1978年）中写道：“个别随机现象虽然是无规律的，但是大量性质相同的随机现象总是有统计规律性的。概率论与数理统计就是一门研究随机现象的数量规律的科学。”这里，我国的数理统计学家们根据哲学的基本观点，明确地指出了“客观大量随机（偶然）现象的规律性”，这在理论上显然是具有极为重要的科学意义的。

可以看到：作为研究统计理论的数理统计学家同在各科学领域研究统计方法应用的社会、经济统计学家们的着眼点是有区别的。前者着眼于数学方面概率分布的抽象大量随机数量规律性的

研究，但根本不问客观的实际情况；而后者则着眼统计方法应用于对客观大量现象数量差异的分析，却不管概率论与客观现象的关系。这种情况显然是由于“认为科学无须求助于哲学”的错误观点所形成的。

综合以上，自1868年凯特勒写《社会物理学》起到现代，为时约一百二十年，列举了许多中外著名的统计学家、经济统计学家的统计学著作。他们对统计学所下的定义和对性质、对象的说明，不仅从统计理论与统计方法不同角度所阐明的重点不同，而且从对统计方法科学研究对象所作的概括也有许多区别。然而，这些统计学家们在对定义的分析中，都提出了统计学的特殊研究对象这样一个根本性问题。就作为科学研究方法的统计学的主要内容来说：首先，统计学是一门特殊的应用数学方法，用以分析研究大量的数量差异（也称作变异）；其次，为了分析研究大量数量差异，统计学从一开始就主要采用平均的方法，来探索分配（或分布）的数量特征；再次，关于大量差异的概率原理，只是在分析样本数量特征的基础上，运用概率分布来推论全域的数量特征，进而作出统计的估计、测验及决策。概而言之，统计学以客观抽象概率为对象的数学方法主要是：（1）大量观察客观差异的统计方法；（2）平均法、标准差、回归相关等分析数量特征、数量关系和变化的统计方法；（3）由大数定律和若干概率分布等所组成的就样本推论全域的统计方法。总之，统计学乃是一门以概率论为基础，分析客观大量现象差异性、研究数量规律性的方法科学。

根据辩证唯物主义的基本观点看来，如果客观不存在集体大量现象的数量差异，就不需要大量观察（或调查）客观社会、自然的大量现象，也就不存在需要综合平均来探索能代表一般水平的平均值。如果客观不存在需要综合平均的大量现象的数量差异，也就谈不上分析客观数量差异的集中趋势和离中趋势，研究

客观数量分布及其主要特征的统计方法。如果客观不存在集体大量现象的数量差异，也就无从应用概率论，更不可能提出和研究从样本推论全域的统计估计推论问题和统计推论方法。

所以，统计学（统计理论或统计方法）的基本问题，就是这门方法科学与客观规律性之间的关系问题。具体地说：一方面，在客观自然界和社会，由于存在着哪些集体大量现象的数量差异性，如零件长度差异，洪水位差，人的身高、体重，工资，消费水平等差异，具有什么性质的数量规律性，才需要应用特殊的统计方法来进行分析研究。另外一方面，在科学理论上，由于什么样的数量关系作为数学推导的客观前题，才引起应用数学方法来分析随机变数的概率分布，探索抽象的数量规律性。统计学作为一门科学，是根据科学研究对象所具有什么样的特殊矛盾而区分的，因而，统计学的基本问题是个特殊的研究对象问题。客观上存在什么样的集体大量差异数量规律性，才构成应用特殊统计方法的对象；而数学理论又以什么样的特殊矛盾作为应用数学分析研究的对象，从而才发展成为一门研究客观数量规律性的方法科学。这个问题不解决，就难以理解和认识现代统计学何以要以概率论为基础，就难以认识辩证唯物主义对于应用和运用统计方法来探索客观集体大量现象数量规律性的指导意义。

三、统计学的基本方法：大量观察与平均法

根据辩证唯物主义的基本观点，客观存在集体大量差异（或变异）的现象，是产生大量观察统计研究方法的客观基础。例

如：洪水位差、零件长度差异、某类树叶子面积或长度的差异、某种作物单位产量的差异、家庭出生婴儿男女人数比例的差异、企业工人工资差异、家庭收入和支出的差异，毛泽东《在中国共产党全国宣传工作会议上讲话》论中国五百万知识分子对待马克思主义的态度的差异（中间多，两头少）等等。正是由于客观自然界和社会存在着种种集体大量现象的数量差异，才产生了观察、分析研究大量差异现象的大量观察法。

1. 集体现象大量数量差异与国家社会全面现象数量情况的区别

早在十九世纪（1836年）作为“统计学之父”的凯特勒就曾宣称：“这种实验的研究法，不得不诉之大量的现象”，“只有基于概率的计算，才可以发现法则”。他不仅指出“大量现象”，而且提到“概率计算”。一个半世纪以来，苏联数理统计学家格涅坚科在其所著《概率论教程》（1954年）的〈绪论〉中，就肯定了十七世纪一些数学家“相信在大量或然事件的基础上能发生明确的规律性”，提到了“自然科学方面（观察误差的理论、射击论问题、统计学问题——尤其是人口统计问题）的迫切要求，使概率论有继续发展的必要，并且引用更发达的解析工具的必要”；他进而指出：“概率论和别的数学部门相似，是由实际要求发展出来的，它以抽象的形式反映着集体性或然事件所特有的规律。”在这里，格涅坚科虽然先提出“在自然科学方面”，但却重点指出“统计学问题，尤其是人口统计问题”，他并未在自然现象和社会现象之间划出不可逾越的界线。而且还指出：这些客观现象都“反映着集体性或然事件所特有的规律”。他说：“按概率论的理解，或然事件具有一系列的特征，特别是，它们都在集体现象中发生。所谓集体现象（大量现象）应理解为是由大量平

等的事物全域中所发生的。”这里，格涅坚科显然在强调：统计学研究的数量规律性是来自客观集体大量现象所发生的规律性。

然而，就在同一时期（1954年）的苏联统计学界，一方面有格涅坚科的代表作《概率论教程》，另一方面却产生了由中央统计局组织作者们集体编写的一本《统计理论》。该书竟保持着十七、八世纪德国纪录学派（康令和作为“统计”一词创始人的阿罕瓦尔）的基本统计观点：把“国家事实”视为“统计学”。应当指出：苏联《统计理论》（1954年）的内容所阐述的是政府统计工作编制社会经济统计数字的若干“经验”，它的基本统计观点，就是把社会统计数字说成是“科学”，而其“理论”实际上乃是外加的马克思主义哲学和政治经济学的“理论”标签，实质上属于社会科学方面的一种“代替论”。《统计理论》一书，正是预先按1953年苏联中央统计局联合科学院和教育部召开的统计科学会议规定的统计学定义“统计是一门有阶级性的社会科学”，“统计学对社会现象研究的结果反映在数字资料中”所制造出来的一本“统计学”代表作。

苏联《统计理论》也提到统计观察的“大量性”。书中说“统计研究的方法就在于对社会现象进行统计观察（按：实指统计调查）”，“统计观察的基本任务，就在于正确地登记个别的社会生活现象（按：实指个别家庭、企业等的情况），这些个别的社会生活现象，构成国内整个社会现象”（按：实指全国及各地区的人口数字、农业生产数字、工业生产数字等）；又说：“统计学的方法决定于被研究的社会现象的性质。由于应进行研究的社会现象的大量性，因而就产生了组织统计观察的必要性，把这种观察作为取得关于个别现象的资料的特殊办法。”这里，苏联《统计理论》所说的“社会现象的大量性”，实际上指一个国家（或国内各地区）经过全面调查所取得的“大量”的人口数字、产量数字等。这种搜集全国基本数字的“社会现象的大量

性”，是从我国古代“禹贡九州”记录全国人口、土地的材料就早已具有了的“大量性”。政府统计工作所搜集的统计数字历来如此，根本不属于现代统计学中作为科学概念的“大量现象”。

所以，苏联《统计理论》所提到的“社会现象的大量性”，乃是一个国家编制全国(及地区)人口和生产数字的“大量性”，这与苏联数理统计学家格涅坚科在《概率论教程》中提出的“集体大量现象”，完全是两回事。苏联《统计理论》所规定的“统计学是社会科学”，实际上是全国（及地区）社会（人口和生产）的统计数字资料，而苏联中央统计局之所以要对全国及各地区编制社会、经济、生产、资源的统计数字，则是为了满足国家对全国进行政治经济管理的需要。苏联数理统计学家格涅坚科写的《概率论教程》所阐明的统计理论，是以抽象的形式研究集体或然事件所特有的规律。统计学是一门科学，作为科学的统计方法是为了满足科学研究、科学实验、科学管理的需要。社会、自然集体大量现象所表现的种种数量差异乃是统计科学研究需要进行大量观察的客观基础。

2. 大量观察

苏联数理统计学家格涅坚科所著《概率论教程》一书，在谈到“数理统计学的基本问题”时写道：“统计研究乃由下列三部分所组成：（1）统计资料的搜集，即搜集表示任何大量全域中各单位的特征的数字资料；（2）所得资料的统计研究，即阐明那些可以在大量观察资料基础上建立起来的规律；（3）统计观察方法的拟定和统计资料的分析，这一部分实际上就是数理统计学的内容。”由此可见，现代社会、自然科学研究应用数理统计方法于分析客观集体大量现象的数量规律性，乃是由大量观察开始的。大量观察是应用统计科学方法于客观实际的基本观

点。

从统计学的历史来说，直到十九世纪中叶，作为“统计学之父”的凯特勒通过统计科学研究的实践，认识到作为探索规律性的大量观察方法（也有人称之为大数定律）。凯特勒在解释大量观察方法时，曾经用一比喻来说明。大意是：先用粉笔在粉板上画一条较粗的曲线，然后站在不同的距离来观察这条粉笔线。如果在最近的距离仔细观察这一粗线时，看见一片大量的、毫无秩序、杂乱无章的小白点；从这些个体的小白点，看不出什么规则秩序。如果稍稍离开来注视，就可以看见多数小白点规则地散布成曲线形态。如果离开粉板远些，远远看去，就看不到个别的小白点，而只观察到集体（或集团）小白点所显示的一条白线的规则性。点与粗线的关系体现了每个个体与大数量个体所组成的集体之间的关系。研究个体的大量所构成的集体时，个体性质即可消失，却表现出集体大量现象的规则性。凯特勒当时就是用这样的比喻来阐明用于统计科学研究的大量观察法的。

二十世纪三十年代曾受德国、日本应用统计学及德国社会统计派观点影响的中国统计学家刘鸿万，在其所编著的《统计学纲要》（1935年）一书的〈统计方法的本质——集团研究〉和〈统计方法的根据——大数法则〉中写道：“欲探求现象的规则性，必须先看在现象中这种规律性是怎样发生的。我们知道宇宙间一切现象的发生、存在、消灭，都是它以前的某种事物状态的结果，一定的结果必由一定的原因而生，现象间的这个关系，便称为‘因果关系’。无论哪种现象，社会或自然现象，都受这个关系的支配，无论哪种现象，只有由这个关系才能说明。不过各种现象虽然都是由一定的原因而发生，但是原因和结果间的关系并不是单一的。为产生一定的结果，必有多种多样的条件、情形、状态结合起来形成一团原因而作用着。这一团原因之中，由对这个结果的因果关系看来，有的是基本的、要素的原因。如果将这

部分的原因除掉，则这个结果便将发生本质的变更。有的是表面的、附随的原因，这部分的原因，无论其有无与否，对于结果都不发生本质的影响。前者称为‘恒常原因’，后者称为‘偶然原因’。因为恒常原因所表现的因果关系是一般的、恒常的、必然的，所以欲产生同种结果，必须有同种的恒常原因存在；有同样的这部分原因存在时，也必生同样的结果。偶然原因所表现的因果关系是个别的、可变的、偶然的，所以欲生同种现象，不必定须这种原因存在；即使它被其它的条件情形所更换，对于结果亦并不发生本质的影响。这两种原因交互错综地作用到一切现象上去，无论哪种现象都受着这两种原因的影响。”

刘鸿万说：统计研究的对象就是“一个大数的集合体”。而这个集合体是“在同一概念，或共同的性质这一点同一的，等质的。譬如，‘北平居民’是由共具‘住于北平’这个性质的个人集合组成的一个集合体，……这些个人的性别、年龄、职业等依然千差万别，含着无限的个别性”。他随后又进一步指出：“对这种由个体的大数集合而成的集体现象施以数量的研究时，则在个体现象中所不能看到的性质，个体现象中所不能显出的规则性，即可在这个集团现象中显现出来，因为个体现象既然主要受着偶然原因的作用，则这些偶然原因作用在个体上是显现着种种不同方向，不同强度的形态；而经常把这些个体的大数集合起来由数量方面看时，这些不同方向、不同程度的偶然原因作用，即相互消杀，相互减弱，于是在个体上影响微弱的恒常原因作用，此时亦因之而扩大、集中、强化以至显现出来。这样显现出来的恒常原因的作用，不外就是这个集团现象中所存在的规则性。”

根据辩证唯物主义基本观点来看，上列这一分析，一方面说明了客观现象由于“恒常原因”与“偶然原因”影响所形成的差异性；一方面又指出了大量观察就可以考察客观集体现象中存在的规则性。这一点与苏联数理统计学家格涅坚科在《概率论教

程》中提到的统计科学研究“大量全域中各单位的特征”和“统计研究大量观察资料”的看法，基本上是一致的。但是，刘鸿万却存在着统计学方面概念不清楚的问题，第一，混淆了“集体现象大量差异”与“国家社会全面现象”（如前述的“北平居民”的性别、年龄、职业这样的“集团现象”）的概念；第二，他未曾能把大量观察和大数定律与他所阐述的数理统计方法联系起来；第三，不能明确“恒常原因”的性质和含义，更不理解“恒常原因”与“偶然原因”的关系。应该说：这三者都是现代应用统计学家们所不能理解的问题。

为此，要理解现代统计科学（如《概率论教程》）大量观察的统计科学研究，就需要阐明大量观察对象的“集体（集体大量现象）”的概念。集体（或集团，或群体）是客观上出现的大量个体现象的集合体；而对集体现象进行统计分析研究的乃是集合体全部个体单位间的数量差异。例如，若干亩生产的小麦单产，若干家庭育龄夫妇出生的男女婴儿数，若干工人的月工资金额等等，这里若干亩，若干家庭，若干工人等就是作为大量现象的集合体。各个个体虽然都具有特殊性（或个性），个别而各自独立的东西，表现为个体间的数量差异；但是，由于这些个体都具有共同的性质，全部个体在共同性质这一点上却都是同一的、等质的，这种具有共同性质的现象从统计大量观察而言，就称之为“集体现象（或集团现象）”。例如，若干种植小麦的田亩，若干由育龄夫妇生育婴儿的家庭，某一工厂若干工人的工资等等，就属于集体中每一个体所共同具有的共同性质。因此，不具备共同性质的大量个体就不属于统计大量观察的集体大量现象。苏联《统计理论》一书提到的“社会现象的大量性”，不过是在一个国家这一空间的全部人口，全部粮食的大米产量，全部小麦产量等的全面调查。全部人口的性别（质量差别）、年龄（数量差别）构成，并不是统计科学研究的集体大量现象的数量差异。

苏联数理统计学家格涅坚科在《概率论教程》中提到"在大量观察资料的基础上建立起来的规律"的问题，他所指的"大量观察资料"不能是任何空间范围（如国、省）内大数量的统计数字，而必须是为了探索集体数量规律性的大量同质现象数量差异的资料。如同为了探索婴儿出生男女比例的数量规律性，就必须建立在大量观察那些生育婴儿的家庭的群体，而这一集体不过是无限全域下的一个样本。这一大量集体家庭在生育婴儿这一点上是同质的。也就是说：这一家庭群体都具有共同的生育男婴或女婴的"恒常"原因。大量观察脱离了具有同质的集体性，尽管统计数字的数量很大（例如，列出一个国家几亿人口的人口数字），也不可能成为"建立规律"的大量现象的基础资料。

3. 平 均 法

根据辩证唯物主义的基本观点，集体性大量差异是统计大量观察的客观基础。统计观察客观种种大量现象的数量差异（或变异）是通过调查（或观察）某种同质的大量个体现象，分别计量每一个现象某一特征的差异数量。例如，统计研究一定阶层（如社员）居民家庭的年收入这一经济特征，首先就要对一定空间、时间下大量同质的（同为社员）家庭进行调查，然后通过调查，取得经过科学计量的大量个体居民家庭的不同收入的差异数字。这样集体性大量个体某一特征的差异数量（如家庭年收入），就是统计观察集体现象大量个体的统计数字资料。从统计研究来说，大量观察的集体数量差异（或变异），不过是统计观察客观现象所分析大量现象数量差异的原料。要进行统计分析研究，还需要对反映客观集体大量的个体差异，进行统计分析（或称为"统计描述"）。如：分析大量数量差异的数量特征，探索某些现象数量规律性的数值，估计多种数量差异之间的数量关系等

等。

统计分析集体大量现象个体具有的差异数字时，第一步就是通过统计整理、分组计数、组织成为次数（或频数）分配（或分布），描述集体大量差异（或变异）分配（或分布）的规则性（也即大量现象所表现的数量规律性）。例如，社员家庭年收入数字资料反映社员集体大量个体（设为1000家）的差异数字，经过分组计数，就组成为按收入组距5元（或10元、20元等）分组的家庭数分配（或分布）的资料。家庭收入分配资料反映一定空间、时间下的家庭集体收入分布的数量（家数）规则性。现代统计科学方法的统计描述部分，就是从反映集体大量现象的次数分配法开始的。

统计描述的第二步就是分析次数分配的集中趋势和离中趋势。统计科学研究之所以要在描述了次数分配的基础上分析集中趋势，是因为客观集体大量现象所表现的数量差异比较普遍的呈现出“两头小、中间大”具有集中趋势的规则性。例如，毛泽东《在中国共产党全国宣传工作会议上的讲话》中指出了中国当时的五百万知识分子对待马克思主义的态度的状况时，就作了“两头小，中间大”的科学分析。这实际上等于说：知识分子的对待马克思主义的态度是有差异的，如果把差异计量下来组成次数分配，也会具有一个明显的集中趋势。（对于马克思主义的态度，不是不能计量，而是目下还没有研究出来科学的计量方法。）既然，在大量观察下的客观集体大量数量差异，在多数情况下，呈现出集中的趋势，作为统计科学研究，首先就要分析这一集中趋势，也就是用一具有分析意义的数值来综合概括集中趋势的典型水平。测定集中趋势的统计方法称之为平均法（平均数有三类：算术平均数、中位数和众数，这里主要指算术平均数）。

英国统计学家鲍来在其《统计学原理》（1901年）的〈平均数〉一章，一开始就写道：“本书为专门研究统计学的，对于平

均数自当予以充分的地位。因为只有应用平均数于复杂的集体（或群体）和极大的数量，才可以以一两个有效数字来予以表明。而且，只有平均数，才可以使统计学的两种定义：平均数科学和大数科学，二者符合一致。”可见，在大量观察的前题下，算术平均的方法对于分析描述集体大量现象数量差异，是具有特别重要的意义的。

奥国应用统计学家齐泽克在1913年写了一本《统计平均数》的著作。他对“平均数”的主要论点，可以概述如下：

他首先说：“差不多各种各类的统计调查，平均数往往用在各种最重要的目的上，因而，可以认为平均数是统计方法中最重要的工具。例如，寿命的平均长度，平均每一家庭生育的儿童数，莱克思（德国，1837—1914）、皮尔生（英国，1857—1936）、高尔顿（英国，1822—1911）等数学家对概率计算的研究等等，都是与平均法有关的统计方法问题。”

他继而在研究到对客观集体大量现象的数量差异，应用算术平均法来计算平均数时，提出两种平均数的问题。他说：“两种平均数，一为客观平均数，一为主观平均数，研究二者的差别，就要联系到两种数列（按：指两种集体大量现象的差异数列）所具有的数值对其平均数的离差。对同一客体所重复计量的数量差异数列（按：指对天文、大地等测量中对同一客体的重复计量）完全决定于计量工具的准确性和平均数计算的准确性，计量的数值或大于平均数，或小于平均数。在这两个限度（按：指正数离差的限度，或负数离差的限度）之内，一定比例的计量数值落在两个限度之中，这可以视为与工具准确性成反比的指数。而计量大量个体不同的主体（按：指如人体测量不同的个人，居民收入调查不同的家庭等）取得数量差异的数列，其离差则起着更为实在的作用。这是因为这种数列的每一个体项目应该视为由于种种特殊原因所决定的事实，而数列所表现的乃是所研究的客观现象的数

量差异。”他接着说："主体平均数与客体平均数的区别是很重要的。如果我们研究统计方法的理论基础，则必须熟悉数学原理，而这些数学原理则是从研究同一客体的重复观察所发展起来的，但是，在现在也应用到'主体'平均数方面。以后者来看，我们的统计研究在应用这些原理于社会学、生物学、气象学等领域相似的不同个体的计量数列时，应该注意在解释上所需要的更改（按：指对两种数列不同差异数量的不同解释）。”

齐泽克在《统计平均数》一书中对平均法的应用提出的论点是很重要的：第一，他区别了客观集体性大量数量差异的两个类型，一种是重复计量的同一客体所形成的大量数量差异，一种是具有种种特殊原因的不同个体（从每一个体本身而言为主体）所体现的大量数量差异。前者应用了平均法，平均数值将接近于客体的真实数值；后者应用了平均法，抵消了种种特殊原因所产生的差异，平均数值将更能代表由共同原因所形成的集中趋势。第二，他指出了统计方法数学原理的客观前提——对同一客观的重复观察——是就大量重复观察的计量差异进行了概率分析，所发展起来的；而这样发展起来的数理统计原理早已有效地应用于社会学、生物学、气象学等领域的不同单位相似个体的计量数列，也即另一类集体大量现象的数量差异（或变异）资料。应当着重指出：平均数计算方法是一样的，对差异的解释是不同的。

现在，再联系前列中国统计学家刘鸿万在其《统计学纲要》中所讲到的大量观察法的应用。他曾提出了现象发生的原因，不外是"恒常原因"和"偶然原因"。他用这样的原因分类来分析"个体的大数集合体"时，说道："由数量方面看时，这些不同方向、不同强度的偶然原因作用，即相互消杀，相互减弱，于是在个体上影响微弱的恒常原因作用，此时亦因之而扩大、集中、强化以至显现出来。这样显现出来的恒常原因的作用，不外就是这个集团现象中所存在的规则性。”他还说：统计方法"便是将某

同一概念下的——即具有某种共同性质的现象个体的多数总括起来，看作一个大数的集合体，对这个集合体由数量方面加以研究，以求具有这个概念的现象中所存在的恒常原因作用。这种研究方法便叫作统计方法。”根据辩证唯物主义的基本观点，应当着重指出：他这样的分析只对了一半，不足之处是没有把大量观察与平均法紧密地联系起来。统计研究方法正是通过大量观察，进而加以综合平均，应用算术平均法使各个个体的偶然性、个别性相互抵消，从而求得具有这个概念的现象中所存在的恒常原因（按：应指必然原因）作用的典型水平，或所体现的集中趋势。仅以这里的“恒常原因”与“偶然原因”的分析为例，一定地区社员家庭平均年收入水平体现了“农村生产形势好转”这一共同的“恒常原因”的影响，每个社员家庭年收入与平均年收入的离差（正数，零数，负数）乃是由于各个差异形态的偶然原因的影响所形成的。

最后，再看看美国经济统计学家密尔士在其《统计方法》（1955年）一书中对于次数分配法、平均法和标准差的阐述。他说：“通过分组，我们可以表明资料的基本结构，揭露集体性大量资料的整体性。为了某些目的，必须把次数分配本身加以概括和简缩，并把分配的性质提炼为三、四个有意义的数字（按：指平均数、标准差、偏态、峰态）”；“用平均数，一个单个的典型数值来代表一个次数分配之所以是合理的，是因为大量数值有集中于中心数值的趋势，而一切观察数值则系规则而匀整地从中心数值分散开来。正是由于在尺度上的事项集中在中心点的四周，这种平均数的代表数字才有意义。”他又说：“实际上任何集团的数字资料，不论社会、生物或经济领域的度量资料，都有差异，都具有单个项目之间数量差别的特征”；“我们现在的目的是要描述观察的次数分配，因而对度量差异的需要将着重于用它来补充中心趋势的度量，而标准差是运用最广泛的而且是标准的

度量差异数的数字。”密尔士还强调指出：“为了确定常态误差曲线下的面积，标准差是习惯采用的单位，因此，标准差在选样和统计推论中具有很大的实际效用。”“全部统计方法确实可以看成为研究差异的一套方法技术。”

由此可见，客观社会和自然界中普遍存在着集体大量现象的数量差异（或变异），这种集体性数量差异，需要人们对之进行统计科学研究，也即采用大量观察和平均的方法。集体大量现象数量差异的普遍存在是产生统计科学的客观基础，而人们应用统计科学这一套方法技术，正是为了探索集体现象大量数量差异的数量规律性。从统计方法的理论发展来看，统计学家们根据大量统计研究实践的经验，早就对影响客观集体现象大量差异的两类原因："恒常原因”与“偶然原因”进行了分析，并有了一定的认识。而关于集体现象大量数量差异的数学理论分析，则是靠数理统计学家们来完成。并随着这一套统计方法技术大量的科学研究实践，从而推动数理统计学的不断发展。

最后，再对大量观察平均的方法从马克思主义哲学的高度，体会恩格斯在分析历史进程中关于必然性与偶然性交互作用下平均数的理论意义，作为本节的小结。恩格斯在1890年9月21—22日给约·布洛赫的信中，首先指出：“根据唯物史观，历史过程中的决定性因素归根到底是现实生活的生产和再生产。……经济状况是基础，但是对历史斗争的进程发生影响并且在许多情况下主要是决定着这一斗争的形式的，还有上层建筑的各种因素”；“这里表现出这一切因素间的交互作用，而在这种交互作用中归根到底是经济运动作为必然的东西通过无穷无尽的偶然事件……向前发展”。接着，恩格斯又进一步分析：“历史是这样创造的：最终的结果总是从许多单个的意志的相互冲突中产生出来的……这样就有无数互相交错的力量，有无数个力的平行四边形，而由此就产生出一个总的结果，即历史事变”；“但是，各个人的

意志……虽然都达不到自己的愿望，而是融合为一个总的平均数，一个总的合力”。①以上就是恩格斯在历史进程的分析中，阐明必然的东西通过大量偶然事件融合成一个总的平均数的论述，应该说：对于统计学平均方法的应用是具有理论指导意义的。

四、统计学的基本理论：大数定律与概率分布

大量观察方法测定了客观集体大量现象的数量差异，而对社会、自然界某些集体大量现象统计调查的资料，经过综合、整理、计算，就可以组成描述客观实际数量差异情况的次数分配（如某时某地大量的社员家庭的年收入分配）。一般地说，实际次数分配的曲线图往往表现为具有集中趋势的规则性。例如，经济统计学家密尔士在其《统计方法》（1955年）中阐明“来自不同领域的次数分配”时，首先说：“宇宙中数量事实都具有共同的法则。”他是在说：宇宙中社会和自然界集体性大量数量差异的各种次数分配，都具有一个共同的集中趋势所体现的数量规律性。他为了证明这一情况，举了许多实际的例子，如：（1）兵士按身高分组的分配；（2）天文测量中观测误差的分配；（3）炮火射击分布区域的次数分配；（4）商品批发价格按年与年之间的变动率分组的次数分配；（5）伦敦对纽约的月份汇兑率的分配；（6）工厂按雇佣工人平均每小时收入分组的次数

① 《马克思恩格斯选集》第4卷，第477、478页。

分配；（7）硬币抛掷试验结果的次数分配。这些分配都具有一个共同的“两头小、中间大”的集中趋势所体现的数量规律性。

根据应用统计学家齐泽克论平均数的两种差异的分类，密尔士所举的许多例子实际上属于两类不同的次数分配，前六个次数分配为大量观察社会和自然领域集体现象数量差异所体现的次数分配；而最后一个属于纯概率领域试验观察的结果，相当于同一客体的重复测量的次数分配。它们在分配集中趋势的形式上是一致的，在性质上却是不同的。前六者是大量观察客观集体现象大量个体数量差异的结果，而后者则根据概率论可以作出数学理论论证的概率分布，这两种数量差异之间究竟存在什么样的关系？它们之间具有一个什么样的共同法则（或规律）？我们将从概率分布与大数定律的理论和实际的发展，作出扼要的说明。

1. 大量观察与大数定律

从早期统计方法的历史来看，十九世纪中叶凯特勒采用概率论：数学理论和大量观察基本统计方法为基础，融合记录（表式）学派、政治算术学派和数理统计学派三者，进行了大量的社会、自然领域的统计科学研究工作，认为大量观察与大数定律（或大数法则）是密切联系的。由于客观现象普遍存在着集体大量数量差异，因而，就用大量观察法来分析研究大量差异的数量规律性（如集中趋势）。并且，由于通过对客观具有偶然性数量差异的大量统计科学研究实践，才认识到在大量观察（包括重复观察）下，会出现几乎必然的大数规律性。

日本受德国记录学派及德国社会学派影响的统计学家横山雅男，在其《统计通论》（1902年）中曾说道：“社会国家的事情，如按科学的大量观察就必然显示大数的法则”；“所谓大数的法则，就是应用科学的大数观察法于同性质的现象，从而使恒

常原因的现象完全显露出来”。他随之又说：“当作出各个个体的观察时，个体中何尝没有恒常原因存在，而所以不能从个体发现，则由于差异的原因掩盖了恒常的原因；例如出生婴儿中，男比女多，虽已成应用统计方法所发现的现象，然而只能在大数中求得，仅仅观察一家甚而一村，也难以显出男女婴儿人数的恒常比例。”这里，横山雅男实际上与德国应用统计学及德国社会统计学派学家的统计观点一样，也在认为：大数定律是大量观察的理论根据。但是，横山雅男不能把实践中所发现的大量观察、大数定律与数理统计方法联系起来，却举了凯特勒就大数定律所进行的一次试验为例。在一壶中放进同数量的黑白球，每次仅随机取出一球，其试验的结果如下：

取出球的次数			球数的百分比	
白	黑	合计	白	黑
1	3	4	25	75
8	8	16	50	50
28	36	64	44	56
125	131	256	49	51
528	496	1024	52	48
2,066	2,030	4,096	50	50

他以此例说明随机取球的次数愈多，黑白球的比例愈近于壶内黑白球数二者相等的真实比例。但是，他却不理解凯特勒关于概率数学计算的思想。

德国应用统计学家苗勒在其所著的《统计理论和方法技术》（1927年）一书中，举过他作过一项在一袋中含有5,000个白球和5,000个黑球真实比例各为$\frac{1}{2}$的科学试验例子。试验结果为：

自10,000球中取出球数	与真实比例的误差%
100	32
200	17
500	9
1,000	8
2,000	4
5,000	0.2

苗勒的《统计学》已经谈到数理统计方法，他有意识地用这一试验数字证明：将随机取出球数继续增加，愈加愈大，则取出的比例与真实比例之差就愈减愈小。

上述两个科学试验的例子，就是应用统计学家们利用黑白球随机抽取的纯概率性质的数量关系，来论证大数定律（大数法则）的。这种数量关系的实质为：观察数愈益增多，则现象出现的比例愈接近现象的真实比例。从应用统计学看来，这种数量关系的规律性就叫做大数定律（或大数法则）。

早期应用统计学家对大数定律是这样解释的：在一定的集体大量个体现象中，每一个体都受“恒常原因”（在上例中，即黑白球全域中各占$\frac{1}{2}$的真实比例的作用），同时，每一个体又具有个别性，也即“偶然原因”（在上例中，即每次抽取球数的随机性，如对真实比例误差的影响）。个体现象的个别性是各个不同的，其特殊作用的方向程度千差万别，互不相同。如果所观察的个体现象数目愈大，则每个个体个别偶然性的不同方向、不同程度的差异，通过平均法，必然相互抵消，其影响的作用愈减愈弱，从而平均数字逐渐显露出“恒常原因”的作用。这一点从大量观察综合平均中，可以看得更为明显。因此，大数定律实质是：观察的数目愈大，则抵消个别偶然原因的作用，而显露出恒常原因作用的可能性愈大。早期应用统计学家正是根据这样的科

学解释，才得出大数定律是大量观察和应用平均法的理论根据。不过，这种理论解释毕竟还是根据试验经验所作出的科学分析。作为统计理论，还需要数学家们从数理统计理论的角度予以论证。

2。大数定律的数学原理

概率论的萌芽状态产生自赌博游戏解决概率问题的需要。十五世纪以来，欧洲意、法、瑞士等国的赌徒们迫切希望掌握赌牌、掷骰、掷硬币等游戏胜负种种可能出现的结果，从而请数学家们寻出支配概率游戏的原理。1654年，法国数学家帕斯卡与另一数学家费尔马相互通讯研究，取得了一般的解法，成为概率原理的基础；1657年，荷兰物理学家兼数学家胡根发表论概率游戏的数学原理，瑞士数学家雅布尔、贝努里于1713年写了一本书专论概率原理；法国数学家达姆瓦于1733年导出一个极为重要的公式，即现代统计学中的常态曲线，也就是作为普遍误差法则的大数定律。由于数学家对赌博游戏中出现概率问题的解法，加以一般化，从而又通过自然和社会统计科学研究的大量实践，就逐渐形成了运用于许多科学领域的一种通用的数理统计方法。

现在，从统计理论的角度，简单地阐明关于大数定律的三个数理统计原理。

（一）二项分布（或分配），也即贝努里所描述的概率分布。

二项分布的概率前提和所发展的概率分布，其要点如下：

设有S种独立事件，每种事件的成功的概率为p，失败的概率为q，而p+q=1，则p与q复合概率可用下式表示：$(p+q)^S$

而各种可能的概率，即展开为：

$$(p+q)^S=p^S+Sp^{S-1}q+\frac{S(S-1)}{2p}p^{S-2}q^2+\cdots\cdots+q^S$$

例如：一个硬币，正面为成功，反面为失败，每一次取两枚

（即S枚）同时上抛落下，进行8次（即N＝8）。在概率理论上，应为：

$$N(p+q)^S；或8(p+q)^2$$

现在，设p和q出现的概率各为$\frac{1}{2}$，则$(p+q)^2=(\frac{1}{2}+\frac{1}{2})^2$；按二项分布展开式的原理，即得：$p^2+2pq+q^2$，而概率分布应为：$\frac{1}{4}+\frac{2}{4}+\frac{1}{4}$。

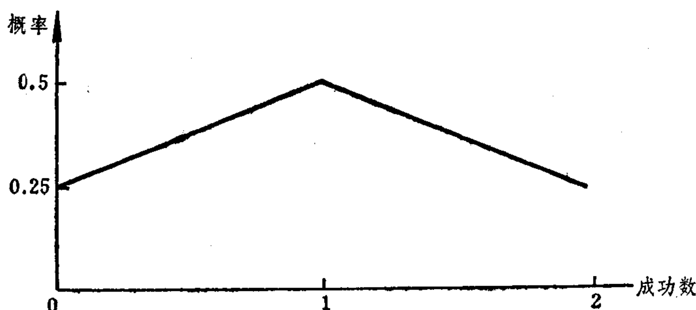
从理论上就是说，出现两个正面的概率为$\frac{1}{4}$，出现一正一反的概率为$\frac{1}{2}$，出现两个反面的概率为$\frac{1}{4}$。这也就是：两个独立事件，正面与反面按排列组合关系出现的各种成功数的概率分布。

现在设试验8次，则按二项分布的原理应为$8(\frac{1}{2}+\frac{1}{2})^2$

抛掷8次（每次两个硬币）的概率分布和次数分配表：

按成功数分组	理论次数	概率分布
2	2	$\frac{1}{4}$
1	4	$\frac{1}{2}$
0	2	$\frac{1}{4}$
合计	8	1

其概率分布的图形为：



对这个“两头小、中间大”的次数分布，也即理论上概率分

布，可以计算平均数。按次数计算的平均数为1，按概率计算的平均数为$\frac{1}{2}$（即两枚硬币中出现1个成功）。理论上p成功的概率分布为0.25。0.5，0.25；概率对其平均数的差异：大于平均数的2。其概率为$\frac{1}{4}$，小于平均数的，其概率也为$\frac{1}{4}$。这种二项式（p和q各$\frac{1}{2}$）的概率分布具有一个以平均数为中心的对称的集中趋势。这个理论上的概率分布说明：

（1）二项分布的成功数成为一系列的数量差异：0，1及2；

（2）各种成功数出现的概率分布（次数分配）体现中间概率高达（$\frac{1}{2}$）和两头概率小为（$\frac{1}{4}$）的数量规律性；

（3）对成功数分组及其概率（次数）的综合平均，求得p的理论（必然）概率为$\frac{1}{2}$（掷两个铜板，平均的成功数为1，就一个铜板言，其理论概率应为$\frac{1}{2}$）。

这里应当着重指出，以上三点与“大量观察客观现象数量差异所反映的中间大两头小的集中趋势；通过综合平均，使差异相互抵消的平均数接近于‘恒常原因’（必然原因）所形成的数值”的实践经验和逻辑推理，基本上是一致的。

美国经济统计学家密尔士在其《统计方法》（1938年）中，曾经举过一个试验的例子。每一次掷12个骰子，1、2、3为失败，4、5、6为成功，试验共掷了4,096次，按二项分布原理计算理论次数。

$4,096(\frac{1}{2}+\frac{1}{2})^{12}$，兹列出试验和理论的次数分布比较如下表：

成功事件	0	1	2	3	4	5	6	7	8	9	10	11	12	总计
实际次数	0	7	60	198	430	731	948	847	536	257	71	11	0	4,096
理论次数	1	12	66	220	495	792	924	792	495	220	66	12	1	4,096

理论与实际基本一致。这一实际例子证明了：作为数学原理

的二项式展开的概率分布，是实验的概率分布的理论根据。

应当说明：这里的概率前提为p等于$\frac{1}{2}$。如p＞$\frac{1}{2}$（或p＜$\frac{1}{2}$），差数愈大，则概率分布线所体现的“两头小、中间大”的规则性就愈不对称。但是，在这种概率的前提下，如果S接近无限大时，概率分布的规则性仍然会表现为对称的集中趋势曲线。这一情况也是大数定律在二项概率分布方面的重要表现。

（二）常态分布。

常态分布也即常态误差定律，甚而如英国统计学家鲍来称之为“大数的定律”。常态分布从数字学原理言，导源自二项分布；常态分布在概率理论中，则为统计学的一个最重要的概率分布。

常态分布在统计理论和实际应用方面，一直居于重要的地位。1733年，法国达姆瓦首先创造常态曲线的概念，规定为不连续型的二项分布在极限条件下的连续型分布。这就是说，二项分布公式：$(p+q)^s$，当S无限加大（即投掷硬币接近无限大次）时，一二项多边形线就变成为一对称的圆滑曲线即在统计学中被称作的常态曲线，这样的常态分布实质上是连续型随机变量的概率分布。常态曲线用方程式表示，即：

$$Y=\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{x^2}{2\sigma^2}}$$

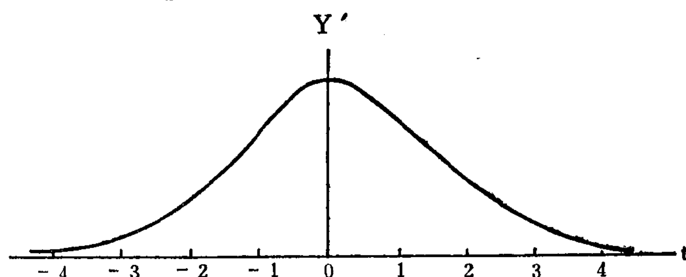
这里，设变数X的平均数为$\overline{X}$，$x=X-\overline{X}$，X分配的标准差为σ。

如再将标准差作为计量单位，以t表示按标准差计量的以$\overline{X}$为中心离差的自变数，即t等于$X-\overline{X}/\sigma$或x/σ。再以Y′代表$\sigma\overline{Y}$。这时，常态曲线的函数就成为：

$$Y'=\frac{1}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}$$

这就是所谓标准的常态分布公式，也即一般统计学书的附表中，以t为自变数，以概率为函数，按积分所列出的常态率曲线

下的面积表。常态分布曲线的纵坐标如下图：



常态分布具有下面一些特征：

（1）绝对值相等的正误差（如t＝＋1）和负误差（如t＝－1）其出现的概率相同；

（2）绝对值小的误差（如 ±1→0）出现的概率大，绝对值大的误差（如2→3，－2→－3）出现的概率小；

（3）绝对值较大的误差（如 ±4）出现的概率接近于0；

（4）由于正负误差的相互抵消，测量中各误差的代数和接近于0，即平均值接近真实的数值。

因此，常态概率分布本身体现了大数定律的数学原理，它是逐步通过实践总结出的大量观察和大量数定律的理论根据。

（三）中心极限定理。

中心极限定理是统计理论中关于随机变数的极限分布的定理，这方面主要有：

（1）“李雅普诺夫”定理。其理论意义是：如果一个随机现象是由很多的随机因素所引起的，而每一因素在总的变化里起的作用非常之小，则可以推断描述这个随机差异现象的随机变数接近于常态分布。

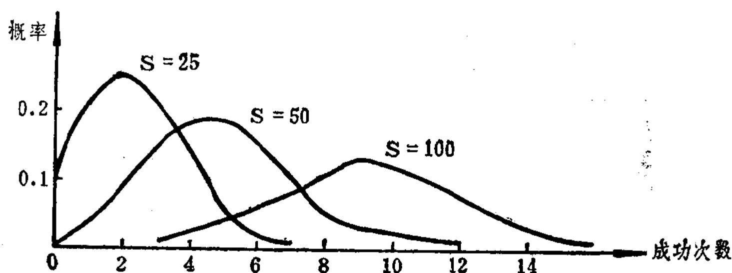
（2）“拉普拉斯”定理。其理论意义是：如果一个随机变数X遵从二项分布，则当$(p+q)^s$的S趋向于无限大时，X趋向于常态分布。

简言之，不论全域为常态分布，还是非常态分布，从抽选的样本分布而言，都受中心极限定理的支配，表现为常态分布的类型。因此，中心极限定理也称作“任意全域的样本平均数的极限分布”。

例如：$(p+q)^s$二项分配，如果$p=0.1$（即$q=0.9$），

设$S=25$，或$S=50$，或$S=100$，

这时，所形成的二项分配曲线，可比较图示如下：



由图看出，当S充分大时，二项分布趋向对称；S = 100时，二项分布接近于常态分布。

从以上有关大数定律的三个数理统计原理来看，前二者，二项分布和常态分布证明：许多属于常态分布的全域，其大量观察的集体数量差异分配是接近常态概率分布的，综合平均，差异相互抵消，必然使得平均值接近于真实数值。后者，中心极限定理证明：统计观察必须大量，前述凯特勒和苗勒的黑白球试验，随着取出球数目的增大，球数比率愈接近于真实比率，或与真实比例的误差愈小。

根据辩证唯物主义的基本观点，应当指出：作为统计理论的数理统计学是应用数学的一个分支，客观现象种种数量关系固然是数学原理推导的客观前提（如赌博游戏对数学家关于胜负机遇的要求），但是，作为研究抽象数量规律性的数学科学（包括应

用数学），则是以一定的抽象数量关系作为推导的前提条件，并在一定假设下加以演绎发展，而求得一些定理、定律及种种公式形态（如常态分布的函数）作为研究的结果的。至于这些定理、定律、公式的导源与客观规律性具有什么样的关系，和它们又如何应用到各个具体科学领域（如物理、化学、生物、社会、经济等科学领域）等等问题，数学科学（包括作为应用数学的数理统计学）本身是不可能、也不应该作为数学问题来加以研究的。这也是“数学是一切科学的工具”的基本特点。因此，各个科学领域在把数学作为方法、技术、工具来应用于具体科学研究时，必须在辩证唯物主义指导下，经过哲学理论思维的科学研究实践，才能取得符合实际的科学结论。各科学领域通过科学研究的实践总结应用数学方法（包括数理统计方法）的经验，反过来又可以促进数学科学（包括数理统计理论）的发展和提高。关于大数定律等概率分布的应用问题，当然也是如此。

3．常态概率分布的应用

常态概率分布是许多概率分布中一个最重要的分布。统计方法的历史说明：通过大量统计科学研究的实践，取得了不少有关常态分布应用的经验。例如，比利时统计学家凯特勒关于“平均人”等等研究，英国生物统计学家高尔顿关于“天才遗传”的研究等。1753年，达姆瓦发现常态分布为二项分布的极限公式后，在一个期间曾被科学家们所遗忘；在十八世纪末叶，为概率论与误差理论的研究者德国数学家高斯和法国数学家拉普拉斯重复发现，特别是认识到大量观察的误差在理论上应属于常态分布的定律（后人往往把常态分布叫作高斯误差曲线）。在人们将常态分布应用于配合和解释客观其它分配的历史过程中，自十九世纪末叶以来，逐渐发现客观现象数量差异的次数分配并不都属于常态

分布，不象高斯等人将常态分布视为一切现象分配的常态定律，经过英国统计学家皮尔生的研究，认为它不过是种种分布中的一种比较常见的类型而已。然而，配合（所谓配合是指按观察样本的次数分配，通过常态函数的配合，来估计理论全域的次数分配）和解译观察的客观次数分配的作用，虽然有一定的限制，但在统计选样和统计估计推论方面的重要性，延至近代，却在日益增加。正是在这个意义上，我们说，在统计理论方面，常态概率分布是许多概率分布中一个最重要的分布。

现在仅从二十世纪以来若干统计学著作中，选择书中所举的例子具体来说明常态分布的应用。

英国统计学家鲍来在其《统计学原理》（1920年）中，举出美国267城市的面粉价格的例子，平均每磅2.629分，而标准差（σ）为0.3334分。对各城平均价格分配的城市数，配合常态曲线的结果如下表：

各城的平均价格分配	实际城市数	曲线配合数
$+3\sigma$以上	2	0.4
$+\frac{5}{2}\sigma—3\sigma$	3	1.3
$+2\sigma—\frac{5}{2}\sigma$	1	4.4
$+\frac{8}{2}\sigma—2\sigma$	9	11.8
$+\sigma—\frac{8}{2}\sigma$	16	24.5
$+\frac{1}{2}\sigma—\sigma$	43	40.0
$0—\frac{1}{2}\sigma$	67	51.1
$-\frac{1}{2}\sigma—\sigma$	47	51.1

续前表

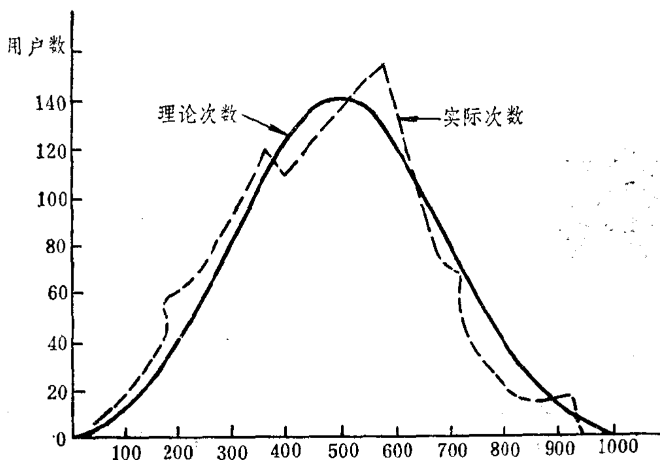
各城的平均价格分配	实际城市数	曲线配合数
$-\sigma — -\frac{1}{2}\sigma$	37	40.0
$-\frac{3}{2}\sigma — -\sigma$	25	24.5
$-2\sigma — -\frac{3}{2}\sigma$	9	11.8
$-\frac{5}{2}\sigma — -2\sigma$	4	4.4
$-\frac{5}{2}\sigma$以下	4	1.7
总　　计	267	267.0

注：按$t = x/\sigma$；$+3\sigma$以上即$t = 3$以上；σ即$t = 1$；$\frac{1}{2}\sigma$即$t = 0.5$等等。

鲍来在这里举了八个例子说明对观察数字配合常态曲线的方法。他在书中写道："为首的二例（如砖的长度）属于直接的测量，表明与大数定律是相符合的；其次二例（如某种族头盖骨及体高的测量）表明在生物测量方面的应用；再一例为学徒年龄与智力，具有间接的关系；最后三例（速度、食物消费及价格）由于现象差异性质复杂而零散，不好估计其全域分配的形式（按：指配合线与实际有某些出入）。"可见，鲍来对于配合问题，主要着眼于曲线的形式，但不能解释随机性模型（常态）与数量差异形式相同的内在原因，更不理解如何进一步分析客观实际数量差异产生的必然和偶然原因。

美国经济统计学家密尔士在其《统计方法》（1955年）中，举美国某地995个电话用户，按使用电话次数分组的用户分配数字，配合常态曲线（如下页图）。

在这个例子中，密尔士用常态曲线配合995个电话用户使用电话的次数分配，所配合的理论曲线与实际曲线是基本一致的。但他对配合有些不大合适的问题只分析了两点："（1）或是由



于样本误差偶然性变动；（2）或是由于观察的分配根本不符合常态误差的定律。”他在这里只看到选样的偶然性误差，看到分配曲线形式的不同，然而却同样的不懂得必须分析客观实际产生数量差异的必然和偶然原因。

苏联数理统计学家格涅坚科在其《概率论教程》（1954年）的〈概率的概念〉一章中论述到：“试验进行次数越大，则观察到的大偏差越稀少。这种频率（指相对比率）稳定性最初是在人口统计方面注意到的。”随之，他就举出了瑞典1935年官方统计资料每月婴儿出生性别分布的例子（见下页图表）。

格涅坚科所举的例子，属于一个在大规模试验中频率几乎不变的统计例证。他指出：“上表表示的是每月生女孩的频率对全年生女孩频率的离差。我们看出频率摆动于0.482这个数的左右。事实上，这频率的波动就围绕着事件的概率p而发生。”现在看来，格涅坚科不象鲍来和密尔士那样，仅知道用常态曲线形式来配合实际观察到的大量现象的数量差异，而是用概率原理来解释

月　份	总　数	男　孩	女　孩	生女频率
1	7,280	3,743	3,537	0.486
2	6,957	3,550	3,407	0.489
3	7,883	4,017	3,866	0.490
4	7,884	4,173	3,711	0.471
5	7,892	4,117	3,775	0.478
6	7,609	3,944	3,665	0.482
7	7,585	3,964	3,601	0.462
8	7,393	3,797	3,596	0.484
9	7,203	3,712	3,491	0.485
10	6,903	3,512	3,391	0.491
11	6,552	3,392	3,160	0.482
12	7,132	3,761	3,371	0.473
全　年	88,273	45,682	42,591	0.4825

人口统计方面出生女婴频率的稳定性，这就为概率论的应用迈出了极为重要的一步。

综合以上关于百年来大数定律（大数法则）的来源、理论和应用来看，它不仅有客观的依据和数学原理的论证，而且通过曲线配合和选样推论，广泛应用于自然、社会的统计科学研究工作实践，证明它是统计方法中的一个重要工具。但是，从数理统计方法应用的角度来看，有一根本问题一直没有得到解决。这个根本问题就是有关概率论数学原理与客观现象集体数量差异的关系问题。历史上，应用统计学家承认客观现象中存在着偶然性，但

不能具体联系到作为应用数学的概率论，最多不过象鲍来和密尔士那样，只看到抽样误差，只能从形式上对待常态曲线的配合是否适合。而数理统计学家按照数学理论和方法，研究和发展了概率论的数学原理，但他们却从一定的概率的数学前提出发，经过推导，最终结束于作为应用数学的概率论，而不能落实到客观集体现象的数量差异是否有偶然原因和必然原因的影响。所以，根据辩证唯物主义的基本观点，就必须研究作为应用数学的数理统计方法与客观集体现象数量差异之间的关系，也就是研究应用数理统计方法，首先应用大数定律这一最重要的概率分布于客观各个实际科学领域的具有指导意义的哲学理论。

最后，还可以引用《马克思主义哲学原理》①一书中所阐明的“统计规律性”与“大数定律”的理论关系，作为本节的小结。《哲学原理》首先提出：“科学上有所谓统计规律性的概念，这种规律性只对各种现象的整体起作用，而不象动态规律性那样对这个整体的每个个别因素起作用。在统计规律性起作用的条件下，大量现象中混乱地散布着某种共同的特征。这种规律性只有通过大量的偶然事件和事实的汇合才表现出来”。《哲学原理》继而又进一步分析说：“统计规律具有规律的一切特征：客观性、必然性、普遍性，它适用于同一类型事物的整体，表现现象的本质的和必然的联系。为了揭示统计规律性，就必须研究大量的偶然性。大数定律是一种表现必然性和偶然性之间的辩证联系的规律。由于这个定律的作用，大量的偶然因素的总合作用必然导致某种不倚赖于个别偶然事件的结果。”这就是辩证唯物主义对于作为统计基本理论的大数定律的马克思主义哲学的科学解释。

---

① 苏联科学院哲学研究所：《马克思主义哲学原理》，第六章〈现实中各种现象的合乎规律的联系〉，人民出版社1959年版，第239页。

五、辩证唯物主义是统计科学应用的理论基础

统计学（也即数理统计学）是一门科学研究方法，到了现代它已成为科学研究、科学试验、科学管理的必要工具。科学研究的目的是为了探索客观现象必然出现的规律性，因而只有在研究客观普遍规律的作为马克思主义哲学的辩证唯物主义的指导下，才能正确地运用现代统计方法这样一套科学研究的工具、方法、技术，来探索客观各个科学领域具体科学研究的数量规律性。

1．必然性与偶然性对立统一的辩证关系

统计理论也是以概率论为基础的数理统计学，是一门研究随机现象（或不肯定现象）数量规律的科学。随机现象在哲学意义上就是由偶然性原因所形成的现象。决定性现象与随机（或不肯定性）现象之间，必然性与偶然性之间，不是绝对分割开来的两种事物，而是对立统一的关系。恩格斯就曾指出："被断定为必然的东西，是由纯粹的偶然性构成的，而所谓偶然的东西，是一种有必然性隐藏在里面的形式"①。

根据辩证唯物主义，我们知道：必然性与偶然性是对立统一的辩证关系。所谓必然性，就是指客观事物发展过程中不可避免

① 恩格斯：《费尔巴哈和德国古典哲学的终结》，《马克思 恩格斯选集》第4卷，第240页。

的、一定要出现的趋向，是由事物的本质原因产生的，是规律性的东西。它是事物发展过程中居于支配地位的趋势，决定着事物发展的前途和方向。所谓偶然性，则是指事物发展过程中可能出现、也可能不出现，可能这样、也可能不这样的现象，是由事物的非本质原因产生的。它对整个事物的发展只起加速或延缓的作用。必然性之所以成为必然，偶然性之所以成为偶然，这并不是由人们随意给它们确定的，而是由于它们产生的原因不同，是由不同性质的矛盾决定的。必然性是由事物的本质的原因产生的，是由事物内在的根本矛盾决定的。可是，必然性和偶然性又是对立统一的：偶然性是必然性的表现，必然性通过偶然性为自己开辟道路。在自然、社会和人类思维中必然性和偶然性不仅同时存在，而且它们是相互依存、相互转化的。必然性是从事物中抽象和概括出来的，是一般事物的本质的东西。它和一般不能离开个别、本质不能离开现象来表现自己一样，它也不能离开偶然性而自己表现出来。任何事物都必取一定的形式，必然性也是如此，偶然性就是它的形式。必然性只有通过偶然性才能表现出来，偶然性是相互依存的另一极。当然，偶然性也不能脱离必然性而独立存在，而是偶然性中隐藏着必然性，必然性通过偶然性表现出来，并为自己开辟道路。不正确地认识和处理必然性和偶然性之间的辩证关系，就不可能正确地认识客观世界。应用统计方法是为了探索客观数量规律性的，应用于分析客观现象的数量关系时，就必须正确地认识和处理必然性与偶然性之间的辩证关系。

数理统计学家专门研究随机现象的抽象数量规律，却把偶然性与必然性绝对地分割开来；应用统计学家专门在某一科学领域应用数理统计方法来探索客观规律性，却不理解概率分布等原理与客观集体大量现象数量差异之间的关系。这恐怕主要是由于主张“科学无须求助于哲学”的实证观点，缺乏辩证唯物主义哲学

理论思维的缘故。

现在就统计观察集体大量现象数量差异的问题，试根据必然性与偶然性的对立统一的关系，来分析一个简单的测量误差的例子。

设测量一块砖的长度，其本身的真实长度为24厘米，并设作了五次测量。测量结果排成序列，成为：

23.8，23.9，24.0，24.1，24.2，平均数为24.0厘米，五个误差：－0.2，－0.1，0，＋0.1，＋0.2；误差之和为0。

个别的五个误差数量属于随机性现象，也就是由那些个别的可能出现也可能不出现的偶然原因（如测量工具摆法的出入，测量者观察角度的差别等）所形成的。测量对象的真实长度24厘米属于决定性现象，也即在测量中五次都居于支配地位的必然原因所决定的。在大量观察综合平均中，偶然性原因所形成的误差，相互抵消，从而显示出接近由必然性原因所决定的真实数值（长度24厘米）。在这里，

$$23.8=24+(-0.2)$$
$$23.9=24+(-0.1)$$
$$24.0=24+(0)$$
$$24.1=24+(0.1)$$
$$24.2=24+(0.2)$$

从现象上看，例如测量数值24.2（或其它四个任何一个数值）这一项偶然性数值就是24必然性的形式；必然性24，只有通过偶然性数值24.2才能表现出来。这就说明了必然性和偶然性不仅同时存在，而且它们是相互依存的。

应当看到：前述奥国应用统计学家齐泽克曾经指出两种平均数和两种数列的问题，两种数列即对同一客体所大量重复计量的数量差异数列（如重复测量一块砖和重复投掷若干硬币等）和计量大量不同个体的数量差异数列（如若干亩的单产，若干工人的

工资，若干家庭的收入和消费等)。这里，区别两种数列显然是很重要的，因为后一数列的离差具有更为实在的作用，为种种具体的特殊原因所决定的事实。因此，要进一步就这后一类大量差异数列说明，同样地可以应用辩证唯物主义必然性与偶然性的对立统一规律来进行分析。如以大量亩数种植谷物（如稻谷、小麦等）的单产为例。共同用一定的种子播在大片一定土质的土壤里，给以一定水份，受到共同的温度，经过一定的耕耘等，这些都是种植某种作物会取得大量面积产量的必然原因。但是大量亩数的单产并不是彼此完全相同的，仍然具有大量的数量差异，共同的必然性通过大量偶然性的数量差异为自己开辟道路。单产之间的差异是受种种特殊、个别的偶然性原因造成的。如种子中某些质量差异，或特殊的土质，水分数量的一些差别，在耕耘中工作质量有别等等。就每一具体亩数的面积看，这些偶然原因是可能出现，也可能不出现的，但是，就某一亩来看，偶然性产量中隐藏着必然性（种子、土壤、水份、温度等），偶然性不能脱离必然性而独立存在。同时，偶然性也是由于偶然原因，如个别种子质量差一些所形成的，又由于有其个别的必然原因，因而，这一亩的单产就低一些。对于这一大片的面积的单产，经过综合平均，大量偶然原因所形成的差异，相互抵消，从而平均单产就表现出接近于种子、土壤、水份、温度等必然原因所决定的亩产收获率。可见，应用必然性与偶然性的辩证关系进行分析是具有实际和理论的重要意义的。

苏联统计学界受记录学派等影响的经济统计学家斯特鲁米林等，在其按苏联《统计理论》观点写的《统计学》（1956年）“统计学的对象和理论基础”中，仍然能按辩证唯物主义基本观点，有分析地说：“个体现象的规模的偶然差异，在个体现象相当大量的情况下，是会互相抵消的。在个体现象相当大量的情况下，这种差异相互抵消的原理，叫做大数定律。概率论所研究的

就是消除偶然差异的影响和表现大量现象的规律性的那些条件。概率论的原理通过数字形式来表现客观的大数定律。”他接着说：“统计学所研究的某些社会现象的平均数，例如：价格的平均水平，平均工资，每一工人的平均产量，都取决于大量个体现象，而个体现象的规模又或多或少地与平均数有差异。在研究这一类现象时，统计学利用的就是大数定律。对统计学来说，大数定律的意义就在于：依靠它的作用有可能在实际可能的范围内，使相应的统计指标脱离偶然的影响。因为由于大数定律的作用，大量过程的总的结果在一定条件下实际已不受偶然的影响。同时，大数定律无论对社会现象的发展水平或发展动态，都绝不起决定作用。它只能使偶然差异相互抵消，从而通过大量社会现象来发现客观的社会规律的作用。”苏联统计学家斯特鲁米林等所著《统计学》，虽然是按照苏联《统计理论》（也即将“社会经济统计数字”视为“统计学的‘理论’”）编写的，但上述论点和分析不仅是根据数理统计学的正确观点，而且符合辩证唯物主义的哲学思想的。

根据辩证唯物主义的观点：必然性和偶然性的对立不是绝对的，而是相对的，因此，在一定条件下它们是可以互相转化的，偶然可以转化为必然，必然也可以转化为偶然。必然与偶然的这种辩证关系，在很早以前就已经由达尔文“从最广泛地存在着的偶然性基础出发”①所创立的物种进化论证明了。例如，前面谈到二项分布中一个 $(p+q)^2=p^2+2pq+q^2$（这里是就掷两个硬币，p和q各为$\frac{1}{2}$而言）的例子。$p=\frac{1}{2}$（或$q=\frac{1}{2}$）为确实概率。这个概率是属于由必然原因所决定的。但在掷两个硬币的条件下，p与q两种概率通过排列组合产生了三种可能性（即p^2，pq，

① 恩格斯：《自然辩证法》，《马克思恩格斯选集》第3卷，第543页。

q^2)。这样，$p=\frac{1}{2}$和$q=\frac{1}{2}$的必然性概率就转化为$p^2=\frac{1}{4}$，$pq=\frac{1}{4}$，$q^2=\frac{1}{4}$的偶然性概率。所以，二项分布（$p+q$）^{2}的展开，就推导出作为按偶然性概率分配的概率分布；而（$\frac{1}{2}+\frac{1}{2}$）^{2}所展开的分配，通过综合平均，抵消偶然性差异，就得到等于必然性概率的$p=\frac{1}{2}$。再如同上列测量误差例，设测量对象为一块砖的长度，则该砖按规格规定的真实长度24厘米为必然性，测量数值24.2厘米、23.9厘米等等为偶然性。但是，如果测量的对象为一千块长度24厘米的砖，分别测量每块砖的长度，得到一千个具有差异的测量数值。这些差异数值都属于具有偶然性的计量数字，只有通过偶然性的相互抵消，求得平均数，才成为接近必然性（制砖规格24厘米）的计量数字。这就再一次证明了，必然性和偶然性的对立不是绝对的，而是相对的，在一定条件下它们是可以相互转化的。

从必然性与偶然性对立与转化来看，前述早期应用统计学家把客观现象产生的原因分为"恒常原因"和"偶然原因"是不确切的，而"固定不变的原因"的提法，则是错误的。现代数理统计学家们把概率与随机现象完全等同起来，似乎是把客观现象分成两类现象，一类是必然现象，一类是偶然现象。这种观点实际上属于形而上学的片面观点，至于受记录学派影响的苏联《统计理论》的观点，把客观现象按照是否具有阶级性划分为自然现象（机械、物理、化学、生物）和社会现象，并把二者对立起来，认为自然现象中才有随机现象，而社会现象中就没有（或很少有）随机现象，显然更是错误的。

毛泽东同志在《同音乐工作者的谈话》中指出："一棵树的叶子，看上去是大体相同，但仔细一看，每一片叶子都有不同。有共性，也有个性，有相同的方面，也有相异的方面。这是自然法则，也是马克思主义的法则。"① 作为学习研究，看来可以这

① 见1979年9月9日《人民日报》。

样理解：这里“共性”就是指“必然性”的具体表现，“个性”就是指“偶然性”的具体表现；“马克思主义的法则”就是指作为马克思主义哲学的辩证唯物主义中必然与偶然或一般与个别对立统一的规律。可以说，客观集体大量现象的数量差异，乃是必然性与偶然性，或共性与个性的对立统一规律的具体表现。作为马列主义哲学的辩证唯物主义对统计方法（也即数理统计方法）的应用，显然是具有理论上的指导意义的。

2. 统计理论与统计方法的辩证关系

客观现象也即自然、社会现象以及人类思维中普遍存在着必然性和偶然性对立统一的规律和必然原因和偶然原因所形成的集体大量现象的数量差异性。这样的客观实际就决定了现代科学研究、科学实验、科学管理需要以大量观察为前提，以概率论为基础的统计学作为一种必要的方法、技术、工具。

根据辩证唯物主义的基本观点，现代统计科学（也即数理统计学）按照科学理论与科学方法的关系分为统计理论（即数理统计学）与统计方法（即数理统计方法）相互矛盾的两个方面。统计理论属于矛盾的主要方面，是由统计理论规定着统计科学的性质；而统计方法则属于矛盾的次要方面，它必须以统计理论为依据。而且，统计理论这一主要方面与统计方法这一非主要方面会在一定的情况下相互转化的。如在应用统计方法的实践中发生了需要有待解决统计方法的问题，这时统计方法（如小样本的分布问题）就暂时成为矛盾的主要方面，提出新的统计方法理论问题，要求统计科学运用数学理论加以研究解决。正是由于统计理论与统计方法矛盾的两个方面在一定情况下的相互转化，从而才不断地促进了统计科学理论也即数理统计理论的发展。

数学是研究抽象数量规律性的科学。统计理论（也即数理统

计学）则是以概率论为基础的一门应用数学。由于客观世界存在着偶然性与必然性对立统一的这一对特殊矛盾，因而，统计科学所研究的特殊对象就是大量现象中个别偶然性与集体的必然性之间对立统一的数量规律性。这里，可以就前面二项分布和常态分布与集体的五个个体测量误差数值为例，来分析作为统计学对象的这一特殊矛盾的数量关系。二项分布例中作为必然性的p和q（例设p和q的确实概率均为$\frac{1}{2}$），它在二项式展开的各项（S＝2，只有3项；S＝100，则有101项），即集体的（S＋1）个项目中，都是存在着的，而作为集体（S＋1）个个别项目分别属于具有种种概率偶然性排列组合的pq。集体的p和q与种种pq排列组合之间的关系，就属于大量现象中集体的必然性与个别的偶然性之间对立统一的关系。测量误差的五个测量数值也是如此，五个数值中共有的2.00属于必然性；五个个别的个体数值如2.02，1.98等就属于偶然性，同样地体现了必然性与偶然性的对立统一。作为应用数学分支的统计理论（也即数理统计的理论）的研究对象就是必然性与偶然性之间这样一对特殊矛盾的抽象数量规律性。二项分布和常态分布正是早期数学家在统计理论方面探索概率数量规律的研究成果。

二十世纪以来，数理统计学家在研究概率分布方面作出许多贡献，如t分布，x^2分布，Z和F分布等等，都属于探索概率数量规律方面的重要成果。应当指出：数理统计学家是以客观集体随机现象作为研究抽象数量规律的客观前提的，因而，只是从数学原理方面推导大量随机现象（即大量抽象的数量差异）的抽象数学规律性。他们自发地看到了确实概率(具有必然性的概率)，但是却未曾认识到辩证唯物主义普遍、一般的客观规律，从而也就不能理解到必然性与偶然性的辩证关系。

数学研究抽象的数量规律，这是数学理论；把研究成果的抽象数量规律作为一定的抽象数量关系，用在处理和分析客观现象

的数量资料，就成为数学方法。统计学也是如此。所不同者，它所研究的是集体大量随机现象的抽象数量差异所具有的抽象的数量关系，因而，把作为统计理论研究成果的集体大量的抽象数量规律，如二项分布、常态分布等应用于处理、分析客观各科学领域（如社会、经济、教育、心理、生物等等）的集体大量数量差异的统计数字时，抽象的数量关系（如前述的标准的常态分布公式）就成为分析客观集体大量现象具体的数量特征、数量关系、数量变化的统计方法。正是在这个意义上，统计理论与统计方法是一对矛盾的两个方面。统计方法属于矛盾的次要方面，而统计理论则是统计方法的根据。统计理论研究的特殊对象是大量现象中个别偶然性与集体的必然性之间对立统一的特殊数量规律。从统计方法来说，则没有什么研究对象，而只有应用的对象，也就是应用于分析、研究客观各个科学领域的某种集体大量现象的数量差异及其数量关系，探索某种具体的数量规律性。

从作为统计理论的现代数理统计学来说，一般地都认为数理统计学是研究随机现象（或随机变数）的抽象数量规律的。但是，有些应用数理统计方法的人，往往认为：这里的随机现象只能是从科学研究工作客观实际随机抽样所取得的样本数值；把随机抽取的数值视为随机变量，从而以为数理统计方法只能应用到随机抽样的样本统计数字（如样本平均数）这类对象上。应当指出：这种看法主要来自一种“社会主义制度下，社会经济现象中不存在随机现象（实指偶然性现象）”的错误观点。实际上，如果不存在客观集体大量现象的数量差异，就不可能产生随机抽样内部和样本之间的数量差异。所以，应用统计方法，首先必须看到：客观自然、社会现象及人的思维中大量数量差异的客观存在。由于客观存在着集体大量现象的数量差异，才决定需要应用以大量观察为根据的统计方法。

应当看到：从统计理论来看，概率分布是关于随机现象，也

即偶然性原因所形成的现象的概率分布。然而，就统计方法应用而言，目的却在于通过偶然性探索必然性的数量规律性。统计研究通过大量观察应用平均法的目的，就是为了测定、估计某一集体大量现象数量差异中由必然性因素所决定的数量特征。例如，观察和计算大量社员家庭收入的平均数，就是为了测定这些集体大量家庭在当前公社，主要是农业生产力水平下所决定的平均每一家庭年收入这一数量特征。大量个体社员家庭收入是具有受种种偶然性原因影响而形成的收入之间的差异。由于这些偶然性的个体收入差异的分配基本上符合常态分布，从而，通过综合平均大量偶然性所形成的差异，相互抵消，平均收入就会接近当前公社农业生产力水平所决定的真实收入水平。再例如，工业产品的统计质量控制，就是在某种产品生产过程中，应用统计方法（包括概率分布）来区分偶然原因所形成的产品质量数量差异与新的（原来生产技术措施中没有的）必然原因所形成的产品质量数量差异。如果发现了新的质量数量差异，就立即停止生产，研究新的破坏质量规格的必然原因；从而提出相应的技术措施，消除新的必然因素，达到控制产品质量的目的。所以，统计理论研究的是偶然性数量规律性，而统计方法的应用，则为了探索、发现、测定、估计必然性的数量规律性。

应当指出：统计理论固然是统计方法的根据，但是统计方法的应用就是以概率论为基础的应用数学方法理论的实践。由于实践是检验统计科学的唯一标准，因而在实践的意义上，统计方法的应用就更为重要。通过统计方法应用的大量实践，一方面实践检验了数理统计学中按一定假设所推导出来的抽象数量规律是否用之有效，起到什么样的实际作用。例如，皮尔生按一定假设所推导的十三个类型的概率分布，虽然推导严谨，符合数学理论，但是，通过大量实践仍然证明不了若干类型是适用有效，符合客观实际的。另一方面，在应用统计方法的实践中，还会发现需要

有待解决的统计方法理论问题（如小样本分布问题），要求统计科学的数学理论加以研究，从而才不断地促进统计理论的发展。而且，数学研究的是抽象的数量规律性，如果不去在实践中应用数学方法，就起不到数学为一切科学探索客观数量规律性的作用。统计理论也是如此。如果没有把统计方法在科学研究实践中应用，就会积存在应用数学理论的“库存”之中，根本起不到统计方法为科学研究、科学实验、科学管理必要工具的作用。

作为统计方法的应用统计学是有应用对象的，具体的应用对象就是：社会、经济、生产、教育、生物、化学、物理等科学研究领域中客观存在着的集体大量现象的数量差异性。从应用统计方法的早期历史看，最重要的统计分析方法就是描述大量现象数量差异性集中趋势的平均法（指算术平均法）和描述离散趋势的方差法（或标准差）。应该指出，由于统计学家在哲学理论上不可能解释概率论与客观大量现象数量差异性之间的理论关系，因而，只能就次数分配曲线集中和离散的形式，来解释应用平均法和方差法的实际意义。其实，如果以辩证唯物主义作为统计学应用的理论基础，就可以认识到：客观集体大量现象数理差异性乃是必然性与偶然性对立统一规律的具体表现。从而就可以进一步理解到：平均法所计量的，乃是必然性（或共性）集中的数量特征；而方差（包括标准差）所测定的，则为偶然性（或个性）离散的数量特征。应该说：平均法是统计描述的基本工具，方差法则为统计推论的基本工具。是以，应用统计学实际上是一套以平均法和方差法作为基本工具的分析必然性与偶然性、描述和推论数量规律性的方法、技术。

3. 统计研究中统计资料与统计学的关系

根据辩证唯物主义的基本观点，理论来源于实践，又服务于

实践，并在实践中接受检验和发展。因此，应用统计方法来分析客观现象的数量特征、数量关系、数量变化，就必须从实际出发，用以探索客观大量现象的数量规律性。各个科学领域存在着共性，具有共同的一般规律，如辩证唯物主义的普遍规律，各领域还存在着许多带有个性的具体的特殊规律。因此，应用统计方法来分析、探索某一具体科学领域的具体数量规律时，就必须从所应用对象的实际情况出发，来考虑各个科学领域运动形式的特点。例如，从生物、社会、经济等科学领域中一些集体大量现象数量差异的样本次数分配来看，生物现象基本上与常态分布相符合；社会现象多数符合，有些有出入（即曲线不甚对称）；经济现象虽有集中趋势，但多数有些出入。因此，从配合常态曲线而言，就需要对经济领域某一集体大量现象数量差异，例如某项产品的价格波动，对具体的经济情况进行具体分析，研究造成价格波动的必然原因和种种可能的偶然原因，进而再大量调查、综合分析、计算测定该项价格波动的规律性。

根据辩证唯物主义的基本观点，作为科学概念，统计资料与统计学是两回事。但从应用统计方法来分析客观集体大量现象的数量差异来看，反映客观具体的大量现象数量差别的统计数字，乃是应用统计方法进行分析的对象，二者组成了统计科学研究一个完整的科学工作，又成为一回事。因此，应用统计方法于调查（观察）分析研究各个领域某一集体大量现象的数量特征、数量关系和数量变化时，就必须按照唯物主义观点贯彻于统计科学研究的全部工作程序，最终达到探索某一科学领域内某一具体的数量规律性（由必然原因所决定的数量规律性）的目的。我们知道，作为统计方法的现代统计学一般地分为统计描述（包括平均法、方差法、曲线配合、回归和相关等）和统计推论（包括概率原理、常态分布、 t 分布、X^2 分布、 Z 分布和估计推断、假设测验、决策分析等）两大部分，这两部分所组成的全部统计方法，有

如现代许多统计学家们作为定义的说法那样："统计学是以种种数学和概率的概念为根据的一套方法"。再如前列《统计理论与统计方法的辩证关系》中所论：应用统计学乃是一套对大量现象数量差异性，分析必然性与偶然性，描述、推论、探索数量规律性的方法、技术。正因为如此，统计学才成为科学研究、科学实验、科学管理的一套工具（或方法、技术）。至于在某项统计科学研究工作中需要运用哪些工具，则要看统计研究的具体问题，例如，作全面调查抑进行随机选样，要分析的是数量特征抑数量关系、抑数量变化等具体数量规律性如何而定。因此，实事求是的统计科学研究方法的全部程序应该包括三大程序：（1）计量描述：调查或观察客观集体大量现象，整理编制如实描述、反映客观实际情况的统计资料；（2）分析描述：即统计描述部分中一系列描述数量特征、数量关系、数量变化的分析方法，如平均法、相关法等；（3）分析推论：即统计推论部分中种种推论、估计、测验、决策等分析方法，也即运用概率原理从样本统计数值推论估计全域真实数值的统计方法。显然，这三大程序是紧密联系的，三者在统计科学研究全部程序中具有一定的内在联系。缺乏对客观全域或样本的集体大量现象数量差异的计量描述，就不可能进行对全域或样本大量观察资料的分析描述（如分析现象的平均特征，分析现象之间的数量关系等）；没有取得经过科学分析的样本数量特征、数量关系、数量变化的数字，就失去进行推论、估计、测验、决策等分析的依据。因此，作为一门统计科学而言，反映客观集体大量现象数量差异的统计资料与统计学是两回事，统计学作为一门方法论科学乃是一套工具；从科学研究的统计研究工作程序来说，统计资料是应用统计方法包括统计描述和统计推论的全部方法的客观根据，因而在统计研究的全部过程中，二者紧密结合，它们又是一回事。按照统计研究问题的需要，科学而有系统地搜集、综合计量取得如实反映客观实际的统

计数学资料，是通过运用统计方法探索客观数量规律性的根据和源泉。而统计方法的理论和一套分析数量关系、变化的工具（方法、技术），则是从客观实际出发，探索客观数量规律性的具体科学方法。所以，根据统计科学研究的具体目的，有计划搜集的统计资料与统计科学是统计科学研究工作中资料与方法的两个方面，二者是密切联系的，又是相互依赖、相互促进的。

应当指出：这里作为应用统计方法探索客观数量规律性的统计数字资料，应该是根据统计科学研究的具体目的有计划地搜集（观察或调查）反映客观某项或多项大量现象数量差异的数字资料。一般目的的统计调查（如普查），其编制的统计资料（如人口普查的统计资料）乃是为了满足行政管理及一般科学研究需要的材料。这种全国及省、市县地区空间分布的数字和质量分类（如国民经济部门分类）的数字，不能反映客观集体大量的数量差异性，达不到区别偶然性与必然性，探索客观具体数量规律性的目的。统计科学研究工作就需要根据研究客观差异性的具体问题，组织专门的调查（或观察）。例如，改革耕作制度，选择优良品种，充分发挥肥效，改善栽培技术，合理密植等一系列科学种田措施，都需要应用统计方法组织田间试验，经过统计分析和推论，探索到具体科学的数量规律性，然后才能择优推广。而为了搞好一项（如作物品种）科学的田间试验，首先又要设计试验方案，固定其它条件（如土壤、施肥等），安排好实地的田间试验，以取得品种不同的收获量，作为应用统计方法（包括概率分布），进行分析的数据（也即科学试验需要的统计数字）。显然，一般大田不同品种的实际单产统计数字，由于偶然因素既多且杂，作为必然因素的播种品种分类既不明确，选种也不严格，是不能用作田间试验进行统计分析推论的统计数字材料的。可见，作为统计科学研究全部工作程序的三个部分：计量描述与分析描述和分析推论乃是密切联系的一个完整的统计科学研究过

程。因而，计量描述的统计数字必须是根据统计科学研究的目的，有计划地搜集能够反映客观集体大量数量差异性的科学数字资料。

统计科学的数学理论是科学的理论，乃是人们通过数学的严密推理，并经过科学实践检验的关于大量现象偶然性与必然性客观数量规律的认识。而统计方法则系以统计理论为根据，用以分析、描述进而估计、推论客观大量现象数量关系的具体科学方法，也是科学研究和科学实验的一套必要工具。我们知道：客观现象的必然性与偶然性是辩证地结合在一起的，偶然性是必然性的表现形式和补充，而必然性又通过偶然性表现自己的作用。据此，在科学研究中应用统计方法，利用概率变化的数量规律的目的，就是为了估计、掌握或控制偶然规律性所形成的数量关系及其界限，来探索、发现或测定必然的数量关系和变化，取得研究有关某一具体规律性的科学数据。综合所述，可以看出：作为马克思主义哲学的辩证唯物主义对于统计科学的数学理论与统计科学方法的实际应用之间的科学关系，对于研究大量随机现象与客观具体规律性之间的科学关系，是具有极为重要的科学理论指导意义的。

六、“实事求是”是应用统计方法的指导思想

毛泽东同志在《改造我们的学习》这一篇马克思主义哲学理论联系实践的论文中，就曾着重指出了马克思列宁主义实事求是的科学态度的问题。他说：“‘实事’就是客观存在着的一切事

物，'是'就是客观事物的内在联系，即规律性，'求'就是我们去研究。”①应该说，“实事求是”这一高度概括的哲学原理，不仅表明了马克思主义的科学态度，而且指出了马克思主义的科学研究方法。从科学态度来说，必须从实际出发，先“实事”，而后“求是”；而不能反过来，从概念、定义出发，先“求是”，而后找“实事”来证明先推理出来的“是”。就科学研究方法而言，“求”就是运用科学的研究方法（包括科学实验、调查研究、统计研究等科学研究的方式和方法），通过多次、反复地实践，来探索作为“是”的客观存在着的规律性。统计学是一门独立的科学，而要应用统计科学的知识，进行统计研究（统计研究指对客观实际大量现象进行数量的分析和研究），就必须遵照“实事求是”的科学态度和科学研究方法。

统计理论与统计方法是统计科学相互矛盾和相互依存的两个方面。统计方法也即数理统计方法，包括作为统计理论研究成果的描述分析方法（如平均法、相关分析法等）与推论分析方法（如常态分布、估计、检验、决策分析等）。是以，现代科学界对于统计方法往往称之为应用统计学或统计分析法。根据辨证唯物主义的观点来看，统计学为应用数学的一个分支，乃是一门方法论性质的科学。统计方法只能具有应用对象（客观存在的社会活动、经济活动、生物活动、化学活动等等客观大量现象的数量情况），而没有象实质性科学那样的研究对象（如客观存在的经济活动、生物活动等）。当然，科学研究者不能直接应用统计方法来分析客观存在的大量现象，而是应用于分析种种反映客观大量现象的统计数字资料。这里的问题是：客观大量现象（或一定空间的全面现象）的任何具体数量方面，都是可以通过统计工作的具体计量而取得反映客观现象的统计数字的。然而，科学研究

---

① 《毛泽东选集》第2卷，第759页。

者对与研究问题有关的一切统计数字是否都能够应用统计方法来探索客观现象的数量特征、数量关系及数量变化呢？换言之，科学研究者是否能够应用统计方法这一套工具，通过任何统计数字来分析、研究、探索客观社会、自然现象中的数量规律性呢？这个问题的中心关键有两个方面：一是反映客观数量情况的统计数字；一是处理、分析统计数字的统计方法。根据辩证唯物主义的观点，是客观数量情况决定是否能够应用某些统计方法，而不是为了应用某一统计方法来挑选一些描述客观存在情况的数字材料。客观现象与方法的关系，就是本节所提出来的：统计方法与其应用对象的关系问题。

1. 两种统计方法与客观现象数量情况的关系

从实际的统计研究工作实践出发，首先应该说明，由于苏联于1954年出版了专门阐述政府统计工作方法经验的《统计理论》一书的影响，统计学界往往分不清楚两种性质不同的“统计方法”。对于数理统计方法称之为“统计方法”，对政府统计工作中搜集统计数字资料的调查方法、整理方法及指标计算方法，也称之为“统计方法”。实际上，这两种“统计方法”是有重要区别的。数理统计方法为“求是”的一种必要工具，是应用来分析反映客观大量现象中必然性的数量规律性的，而政府统计工作方法则指运用调查、整理和指标计算等方法，属于“实事”的重要方法，是用来取得表明全国及各地区某一时期的社会、经济、生产、资源等统计数字资料的。这里，为了避免两种统计方法的混淆，可以用“统计分析方法”（即统计方法的同意语）代替作为数理统计方法的“统计方法”。应该说：反映客观现象统计数字资料是应用统计方法来分析、研究的对象。它们二者乃是统计资

料与统计学（数理统计方法）的关系。

其次，必须指出：从应用数理统计方法分析研究各社会、自然科学领域某一集体大量现象数量差异性，来探索存在的具体数量特征、数量关系、数量变化来看，不仅要评定这一统计研究是否遵照客观数量情况的具体特点，来正确选择数理统计方法中某些适用的分析工具；而且更要检查这一统计研究所使用的种种通过统计工作方法搜集的统计资料，是否能够准确如实地反映客观现象实际存在着的数量情况。显然，任何脱离客观实际的统计数字，或者错误地应用了数理统计方法的某些统计分析工具，都得不到正确的统计结论。从辩证唯物主义的观点来看，客观存在的实际情况乃是第一位的，统计分析方法则为第二位的。统计数字的正确与否，决定于它们是否反映客观现象的具体数量情况；统计方法应用得是否科学，决定于统计方法、这一套统计分析工具的应用是否符合客观存在着的具体的数量关系。对于后者，应该指出：人们的主观认识方法必须符合客观实际，而不应该要客观实际来适应主观意识中希望使用的某种统计方法（工具或方法、技术）的需要。例如，人们为了使用相关分析法，竟按主观愿望将客观上并不具有内在联系的数字资料，进行纯数量上的相关分析。这种应用显然就是错误的。

2．统计分析方法与反映客观现象数量情况的统计数字的关系

实事求是地说，数字与数学方法是两回事。统计数字是反映一定空间和时间条件下客观实际存在的自然和社会中某一具体现象的数量情况的，因而，作为统计资料的许许多多统计数字，都是经过科学的统计工作（政府的统计工作和种种科学研究工作中观测、调查、计算客观数量情况的统计工作）实地计量的单个的

统计数字，如：1982年7月1日零时中国人口10,038万人，长江现长5,800公里。就每一个在具体空间和时间条件下所计量的单个统计数字而言，它只能表明某一具体现象在某一空间、时间条件下的具体数量，本身不具有任何更多的作用（如用以分析数量关系的作用）。正是由于统计数字的这一特殊性质，现代国际科学界的科学研究工作，对许多统计数字所组成的材料往往称之为“资料”。只有各国政府统计工作所搜集编制的社会、经济、生产、资源等系统的统计数字材料才称为“社会、经济、生产、资源统计”，并将每年全国及各地区系统表列出版的分门别类的统计数字的材料，称之为《统计年鉴》。至于各个科学领域科学研究中的统计工作，则如前节所论，仅仅是统计研究工作三大程序（计量描述、分析描述、分析推论）中的第一个程序；其统计材料往往不另行独立地编印，而只是列入研究论文，供本科学领域科学研究的参考。

现在，就以现代各国政府统计工作每年出版的《统计年鉴》中的一般内容为例，阐明对于系统组织的统计数列应用统计分析方法的问题。政府统计工作（和未另行独立编印统计资料的科学研究统计工作）所编制的系统组织的统计数列，也即一系列互相联系的统计数字资料，可以分为四类数列：（1）地区数列，也即按空间地区分类的数列（如各省的人口数，经济地理科学研究工作中若干按地质分类地区的小麦产量等）；（2）质量分类数列（如国民经济各部门的产值，医学研究中死亡病因分类的死亡人数等）；（3）时间数列（如全国每年钢产量的数字，科学研究工作中的某种作物生长按日观测的数字）；（4）数量分组数列（如人口年龄分组，人民生活水平研究工作中的居民家庭收入分配，工业产品质量控制研究工作中的钢丝抗张强度的次数分配等）。应当指出，上列四类统计数列概括了一切系统地反映客观实际存在着的相互联系的具体数量。正是由于四类彼此不同的统

计数列，分别反映着客观存在着的四类不同的具体数量联系，才决定了人们必须采用不同的统计分析方法，进行相应地统计描述分析。如果为样本的统计数列，就需要进一步进行统计推理分析。

然而，四类统计数列表明客观现象中四类不同性质的数量联系，体现为四种不同的分类或分组下的具体数量差别，这就需要运用相应不同的统计分析方法。

（一）空间分类数列。

空间分类数列表明客观上不同地区之间的具体数量差别。例如，全国各省人口分布，有的省人口多些，有的省人口少些，这就首先需要计算各省人口占全国人口的百分比，应用比例法来进行统计分析；同时，还需要运用地图的图示法来研究分布情况。

（二）质量分类数列。

质量分类数列表明客观上若干质量不同的类别间的具体数量差别。例如，国民经济各部门的产值数字，农业部门多些，工业部门少些等等，也只能应用比例法计算各部门产值占整个经济总产值（以国民收入言，即国民生产总值）的百分比，进行统计分析。

（三）时间分组数列。

时间分组数列或动态数列（时间可以按年、季、月、日、时等时间单位进行分组）反映客观上某一具体现象的数量变化。随着时间的自然变动而变化的具体现象，往往包含有数量变化的规则性。例如，一个国家人口在年度分组下的自然增长，往往包含有几何级数的增长变化。作为统计分析，就需要应用统计描述分析的方法（如曲线配合法），而且对于有些时间数列还需要应用统计推论分析的方法（如分析随机残差踪迹的自身相关法）。至于对时间分组数列或动态数列，直接应用比例法计算发展速度，则属于一般的个体指数法，用以简缩数字、加工处理。

（四）数量分组数列。

数量分组数列反映客观存在着的某一集体大量现象的数量差异性。应当特别指出：客观存在的集体大量现象的数量差异性，乃是建立以概率论为基础的数理统计学的客观前提。从应用统计方法来说，在客观自然和社会现象中，集体大量现象所表现的种种数量差异正是统计科学研究需要进行大量观察、综合平均，进而分析集中趋势和离散趋势的客观基础。所以，对于数量分组数列的分析、研究，乃是运用统计学的统计描述分析法和统计推论分析法这一套统计分析工具，按照统计研究需要的具体应用。应当指出，客观种种具体的大量现象的数量差异是必然性与偶然性对立统一规律性的数量表现；只有应用统计分析方法才能通过数量分组数列，分析和研究客观现象中必然性的数量规律性。

3．“实事”，才能应用统计分析方法来“求是”

上列四类统计数列资料的根据是客观存在着的实际情况。但是，统计数列分为四类，并不等于客观存在的社会、自然现象也分为四类现象。四类统计数列乃是统计工作从空间、时间、质量、数量四种不同的角度，对同一对象的各种特征，分别进行大量观察、计量数字的结果。例如，作为社会主体的家庭人口，分散地居住在一定国土之内。统计工作对一个国家的全部家庭人口，按其各种需要研究的特征，从不同的角度，可以计量取得四类不同的统计数列。例如，对于家庭的人口数特征，从空间各地区分别计算人口数，就可以得到全国人口地区分布的统计数列；从每年的同一时点（如7月1日零时）按年连续计量全国人口数，就可以得到按年变动的时间数列。对于家庭人口的民族特征，按照实际的民族类别，分别计算各民族的人口数，就可以得到全国人口按民族差别的质量分类数列。由于人民生活水平的高低是以家庭作为基本单位的，因而对于家庭人口的消费支出特

征，就要按照每一家庭的年（或月）消费支出金额，进行数量分组，求得全国家庭年（或月）消费支出的数量分组数列。

从反映家庭人口对象诸特征的四类统计数列来看，只有时间数列和数量分组数列才可能具有一定的数量规律性。正是由于家庭人口对象存在着由这两种数列所表明的客观情况，才决定了统计研究工作采用相应的统计分析方法。如：对人口的增长变化采用曲线配合法进行统计分析；对于家庭人口的消费支出，就采用组织次数分配等统计描述分析法（如平均法、标准差法等），以及统计推论分析方法（如平均数的估计和检验）。这里，对于用时间数列来描述客观数量变化是否具有普遍性的问题，还需要进一步根据辩证唯物主义的基本观点，来考察家庭人口的社会运动不同于某些自然运动的特点。我们知道：客观世界的五种运动形式：机械、物理、化学、生物、社会，在内容上不仅具有从低级到高级的不同，而且还有无机界和有机界之别。社会运动属于最高级的有机界的运动形式，才呈现出随年代而迅速变化的人口再生产。由于社会人口发展变化中具有一定的数量规律性，才导致统计研究工作对之组织按时间分组的统计数列，并且采用曲线配合等等统计描述分析方法。无机界的机械、物理、化学等较低级的运动形式，就不存在随着年代进展而变化着的规律性。在统计数列方面当然就只需要组织具有数量差异的数量分组数列，而不需要组织研究动态变化的时间分组数列。根据以上有关四类统计数列的举例分析，充分证明了：是客观存在现象的实际情况，决定着是否需要应用统计分析方法，来探索客观对象所固有的数量规律性。

最后，再应用"实事求是"这一马克思主义哲学原理来小结：统计分析方法与其应用对象之间的关系问题。按照"实事求是"作为统计研究方法论的指导思想，显然应该"实事"在先，"求是"在后。而"实事"就是从客观实际情况出发，通过统计

调查、整理、指标计算等统计工作的方法，取得所需要表明的客观实际情况的数字资料，主要是反映大量现象的数量分组以及时间分组的统计数列。“求是”则主要是按照研究客观情况和问题的需要，通过研究，灵活地运用统计描述分析和统计推论分析这一套统计分析方法，尽可能地探索到客观大量现象本身所固有的必然性的数量特征、数量关系及数量变化，也即客观大量现象中的数量规律性。

七、具体地分析统计方法在科学研究客观规律性中的具体作用

列宁说过：“马克思主义的最本质的东西、马克思主义的活的灵魂：具体地分析具体的情况”①。应该说，“具体地分析具体的情况”这一马克思主义研究问题的方法原理，对于一切自然、社会科学领域的科学研究及统计研究，都是适用的。从应用统计分析方法的统计研究来说，它是为了分析各类统计数列的统计资料，研究和探索客观大量现象中属于必然性的数量特征、数量关系及数量变化的。现在为了探讨科学研究与统计方法的关系，就必须根据辩证唯物主义的基本观点，对统计方法在研究客观具体规律性中的具体作用的问题，按照客观存在情况的特点，进行具体地分析。

根据辩证唯物主义的基本观点，宇宙间的一切物质运动都有其反复出现的确定趋势，贯串其中的规律性都是不以人们的意识

① 《列宁选集》第4卷，第290页。

为转移的客观存在。而各种物质运动形式中的具体规律，则系事物发展中本身所固有的本质的、必然的、质量统一的稳定联系。科学研究是为了探索客观存在的规律性的。在科学研究中，确定事物及其状态的性质，叫作定性研究；对事物进行数量分析，则系定量研究。定性是定量的基础，定量是定性的精确化。至于统计学则为一门应用数学，乃是科学研究在大量观察下从事定量分析的一种必要工具。是以，从科学研究方法来看，就需要探讨统计方法这种定量分析的工具适应研究对象的客观基本运动形式和客观各种运动具体规律性的特点。

1. 客观五种运动形式的特点与应用统计方法的关系问题

按照辩证唯物主义的基本观点，客观世界存在着五种基本运动形式，即机械运动、物理运动、化学运动、生物运动及社会运动形式。它们具有从低级到高级的发展顺序，其中机械运动为最低级的，而社会运动则为最高级的。五种运动之间的关系是密切联系，层层包含。物理运动除包括机械运动外，并另包括声、光、电等物理运动形式。化学运动则包括物理运动在内的全部化学运动形式。生物运动包括化学运动在内，而社会运动则包括从机械到生物运动在内的全部社会运动形式。是以，科学的分类，必须以客观物质的普遍联系的各种基本运动形式为基础。相应地，在五种客观运动基础上所建立和发展的社会科学、生物学、化学、物理学、机械学，实质上为五大类密切联系、相互依存的实体性科学。同时，马克思主义哲学也即辩证唯物主义或唯物辩证法，系以客观存在的全部五种运动形式作为研究对象，它既是实体性科学，也是方法论科学。至于数学，它所研究的则系以客观存在数量情况为前提的抽象数量规律性。从应用言，数学也属

于方法论性质的科学。而统计学则为数学分支的一门应用数学，当然，也属于一门方法论科学。以上所论，就是以客观存在的五种运动形式的特点为基础，根据人们的长期研究、认识和实践，发展起来的现代科学的基本类别和它们之间的相互关系。

客观存在着的机械、物理、化学、生物、社会等五种运动形式，它们之间密切联系，相互依存，形成了世界上整体的客观运动。但是从客观存在具体情况的特点而言，这五大类运动形式则有重要的区别。一方面，五种运动有低级和高级之别，另一方面，五种运动之间又有无机和有机之分。后一情况的出现，则是由于在世界历史发展的长期过程中，无机界（机械、物理、化学）的一分为二，在原有无机界之外，因生命出现，又产生了有机界（生物到人类社会）。这样，世界上才出现无机界和有机的五种运动形式。应当指出：这五种客观运动形式之间的这两个特点，对于应用统计分析方法是有重要关系的。

第一，由于五种客观运动形式在内容上具有从低级到高级的差别，这一特点本身就表明了：从影响各种规律性的因素来说，低级运动形式影响因素的数量比较少，因素之间联系比较简单；级别愈高，影响因素愈多，因素之间的联系愈益复杂。社会运动形式是最高级的，各种规律性本身所受的影响因素的数量最多，因素之间的联系也最为复杂。根据“对具体情况进行具体分析”的原理，针对着五种运动形式由低级到高级的重要区别，在指导思想上，就应该对具体的运动形式的具体对象，进行具体分析，然后再研究需要采用的统计方法。例如，对于低级的机械运动形式的统计数列进行的相关分析，就需要采用简单相关分析法；对于高级的社会运动形式者进行的相关分析，就需要运用多元复相关分析法。

第二，无机界的机械、物理、化学等较低级运动形式的种种规律性，在历史的过程中循环往复，并不因为时间的转移而出现

变动。有机界的生物到人类社会等较高级的运动形式，一者由于具有生命继续不断地再生产，就使得在数量上由少到多；一者从植物、动物到人类社会还在不断地变化，就使得在质量上出现新的演化。因之，有机界的种种规律性不仅具有循环往复的变化，而且生物和社会的运动还随着时间的转移而在不断地发展变动。应当指出，机械、物理、化学等无机界的运动形式与有机界的生物到人类社会运动形式之间的这种主要区别，对于统计研究具有重大的关系。有机界的动、植物和人类的自然再生产是随着时间的进展而变动的，而且人类所组成的社会和所从事的经济活动又是随着历史的进展而发展变化的。因此，动、植物和人类社会的演变和发展，都具有一种按时间（时、日、月、年）进展而运动着的规律性。是以，统计研究不仅需要通过统计计量描述方法来观察或调查客观存在的生物和社会的动态变化，整理成为时间分组数列；而且还需要进一步应用统计分析描述方法和统计分析推论方法，探索客观存在着的动、植物界和人类社会某些具体的集体大量现象中的数量规律性。这样，作为应用统计方法的原则应当是：按照客观运动形式的具体特点，采用具体的计量方法。同时，按照五种运动形式的特点，予以区别对待，首先，组织数量分组数列，有条件才可组织时间分组数列；然后对具体的数列（如一个、两个、三个变数数列）运用具体的统计分析描述或统计分析推论方法。

2．统计方法在科学研究探索客观规律性中的特殊作用问题

统计方法是科学研究的一个必要工具。现在在研究了客观五种运动形式特点与应用统计方法的关系的基础上，进一步探讨：科学研究工作应用统计方法这一套统计分析工具，在探索客观规

律性的认识过程中所起的特殊作用问题。首先，在统计研究中，应用统计方法分析客观集体大量现象的数量差异性，主要是为了分析、掌握、控制偶然性，来探索和测定必然性的数量特征（如算术平均数）、数量关系（如回归相关）及数量变化（如动态曲线）。应该看到：这些经过应用统计分析方法所取得的科学数据，进一步再经过多次科学实践的检验，才能达到反映客观数量规律性的水平。不过，即使经过反复实践，成为经过验证了的某一具体的数量规律性，它仍然只是客观存在着的某一客观具体规律内在联系的具体数量表述。当然，关于客观规律的数量表述是很重要的，因为它能具体、准确地表明客观规律性所具有的数量关系。我们知道：人们对客观规律的反复观察（调查）和研究，由感性认识提高到理性认识，从而能够用文字概述人们对某一客观具体规律的认识。这固然是一项科学研究客观规律性所取得的科学成果，但是还不够完善。马克思认为："一种科学只有在成功地运用数学时，才算达到真正完善的地步。"①现代物理学的科学研究充分运用了数学方法，作为科学研究成果的种种物理规律（或定律），几乎都是用数量关系表述的。应该说：对于某一客观规律的数量表述，乃是该规律的内在联系与数量表现的统一。如果统计研究的定量分析只取得关于客观规律的数量规律性的表述，而缺乏关于客观规律内在联系的定性分析，未能透过现象认识本质，显然，更是完全没有达到真正完善的地步的。恩格斯把数学就看成是："辩证的辅助工具和表现方式"②。至于统计学，则只是科学研究在大量观察条件下分析偶然性与必然性，测定并探索必然性的数量规律性的一套工具。统计方法并不能够、也不可能

① 转引自拉法格：《回忆马克思》。

② 恩格斯：《自然辩证法》，《马克思恩格斯选集》第3卷，第521页。

为各门实体性科学（物理、化学、生物、经济学、社会学等）分析研究客观规律性的内在联系。例如，在具体研究经济科学领域中的某一具体规律（如价值规律）时，用来控制偶然性、测定必然性的数量关系的工具，固然是统计学（统计方法）；而从经济理论上具体分析规律的具体内在联系的，则为经济学。

因此，对于统计学与实体性科学研究、探索规律性之间的具体关系问题，不能抽象地阐明统计方法与一般科学研究的关系，而只能对具体科学领域的具体科学研究，进行具体的分析。这里，仅以经济学和人口学中科学研究探索客观具体规律性的问题为例，来阐明应用统计方法与经济学理论或人口学理论之间的具体关系。

（一）经济学方面关于价值规律的分析研究。

这个问题可以马克思在《资本论》中对于市场价格变动规律性的经济分析和统计分析为例，说明经济规律性与其数量规律性之间的关系。马克思指出："市场价格固然会高于或低于这个起调节作用的生产价格，但是这些变动会相互抵消。如果我们考察较长时期的物价表，把商品实际价值因劳动生产力变动而发生变化的情况和生产过程因自然事故或社会事故而受到干扰的情况撇开不说，我们将感到惊奇的是：第一，各次偏离的界限比较狭窄，第二，这各次偏离的平衡具有规律性（按：指价格变动的数量规律性）。在这里，我们也将发现凯特勒在社会现象上论证过的那种起调节作用的平均数的统治作用。"①这里，马克思对于市场价格因劳动生产力变动而变化的分析，乃是根据经济学所作出的经济分析（或经济解释）。而马克思肯定了的凯特勒（即前面提到的"统计学之父"的应用统计学家）对于社会现象所论证的那种起调节作用的平均数的统治作用，则系由于凯特勒应用了概率分布，研

① 马克思：《资本论》第3卷，第972—973页。

究社会现象方面人的生理、精神等数量特征的数量规律性及其平均数（按：即“平均人”），所作出的统计分析。马克思所作出的经济理论分析，乃是关于价值规律中价格变动内在的因果关系；而被马克思肯定的凯特勒所曾经作出的统计分析，则为市场价格变动所表现的数量规律性（属于必然性的数量特征、关系及变化）。就这两种价格分析的科学研究方法来看，实际上是把关于价格变动的经济学分析与统计分析结合在一起。从对价值规律的科学研究目标而言，则是把价格变动的规律性与表述这一规律性的数量规律性统一起来，使价值规律的研究“达到更为完善的地步”。在这里，同时还要进一步看到：马克思在1869年3月3日给路·库格曼的信中，又对凯特勒提出了批评，说他“过去有很大的功绩。他指出，即使是社会生活的表面上的偶然性，由于它们周期性的反复和周期性的平均数，仍旧具有内在的必然性。但是他从来没有能对这一必然性作出解释”①。这里，应当看到：马克思在《资本论》中，肯定了凯特勒运用数理统计方法所作出的关于“平均人”等“平均数的统治作用”的统计分析，而在给路·库格曼的信中，除了肯定外，还批评了凯特勒对社会内在的“必然性”“从来没有能作出”社会经济的科学“解释”。可以说，作为定性分析的社会经济分析与作为定量分析的统计分析，乃是对同一规律性的质与量两个方面不同性质的分析。马克思肯定凯特勒的为定量分析；而批评他的乃是定性分析，也即他在社会经济方面从来没有能作出科学的解释。另外，这一情况也说明了：对客观具体规律性的科学研究，仅仅进行统计分析，探索到一些具体的数量规律性，那是很不够的，必须把定量分析与规律性内在联系的定性分析紧密结合，取得一致。

（二）人口学方面关于出生性比例规律性的分析研究。

---

① 《马克思恩格斯全集》第32卷，第583页。

在人口统计研究的长期实践中，曾经多次通过大量观察，应用了综合平均的方法，探索到人口出生性比例总是表现为52男婴对48女婴这一必然性的数量特征。应当指出：52比48这一数量规律性只是统计分析的结论。根据辩证唯物主义的基本观点，社会运动形式属于包括生物运动形式在内的最高级运动形式。因此，对于人口出生性比例的规律性，还必须深入到人口生理方面，进行生物学的分析研究。只有倚靠生物学的具体分析，才能探索到出生性比例内在的必然联系。这一点，根据当前生物学的科学研究①，在理论上业已取得有关规律性的认识为：（1）人类细胞中有23对染色体，其中有一对为性染色体；而性染色体又分为X染色体和Y染色体，男性的染色体由X与Y组成的XY，女性者则由X与X组成XX。（2）人类的性别就是由此决定的，男性精子细胞的染色体XY型分裂成为X型和Y型两类精子；如果X型精子与X型卵子结合成为XX配对，即为女胎，如果Y型精子与X型卵子结合成为XY配对，即为男胎。（3）精液里所含的X精子和Y精子两类的数量是相等的，而和卵子的结合则是随机的，各有50%的可能性；但实际上Y精子与卵子结合者比X精子与卵子结合要多，因而从整个人类来说，人口出生性比例总是维持在52∶48。以上三点，就是现代生物学对人口出生性比例这一具体的生物规律性所进行的具体的生物学分析及相应的统计分析。这里，从科学研究的角度来看，上列分析研究的第（3）点，是不能解释出生性比例何以是52比48，而不是50比50的问题的。就科学研究方法而论，出现这一问题的原因是：对于生理上精子与卵子结合的具体数量关系，并未通过科学实验或统计估计，取得科学的具体数据。现在竟简单地假设为：“精液里所含的X精子

---

① 见林巧稚等：《家庭育儿百科全书》上册，〈性别的决定主要在精子〉，北京出版社版。

和Y精子两类数量相等”，这在方法上是不科学的。从统计理论来看，人口统计研究的52比48这一具体的数量特征，实际上为长期在多次大量样本观察下，精液里X精子和Y精子二者比例的概率分布的平均数。根据统计科学中的抽样理论，这一经验的平均比例，就是对二项分布中p或q的理论概率，也即必然性比例的估计。只有进一步作了具体的统计分析推论，才算完成科学的定量分析，取得了必要的科学数据。是以，这一人口出生性比例的分析研究是在说明：必须从人口统计研究出发，深入地分析研究到人类的生理活动，在质与量两个方面，把具体的生物学分析与具体的定量分析结合起来，才能对人口出生性比例规律性的特点，作出科学的具体分析和具体解释。

通过上列两个科学研究实例的具体分析，可以看出：实体性科学的科学研究对其具体规律性的定性分析，乃是主要的研究方法，而在探索规律性过程中应用统计方法来进行定量分析，则属于科学研究中的一种必要工具。

最后，可以再探讨一下：在科学研究中应用具体地分析具体情况的方法原理，来探索具体的规律性的问题。前面曾经提到：毛泽东同志对“实事求是”的论述：“‘是’就是客观事物的内部联系，即规律性”。毛泽东同志还在《矛盾论》中指出：“事物发展的根本原因，不是在事物的外部而是在事物的内部，在于事物内部的矛盾性。任何事物内部都有这种矛盾性，因此引起了事物的运动和发展”；“一切事物中包含的矛盾方面的相互依赖和相互斗争，决定一切事物的生命，推动一切事物的发展”①。而且，恩格斯也曾指出：“运动本身就是矛盾”②。可见，根据辩证唯物主义的观点，正是由于客观具体规律性的具体内部联系，才体

① 《毛泽东选集》第1卷，第276、280页。

② 恩格斯：《反杜林论》，《马克思恩格斯选集》第3卷，第160页。

现为特殊的矛盾运动；正是由于事物内部的特殊矛盾性，才引起具体事物的特殊运动和发展。科学研究的目的是为了探索和发现客观存在的某一特殊而具体的矛盾运动规律性。而作为唯物辩证法方法原理之一的“对于具体的事物作具体的分析”，主要是分析、研究具体对象内部相互联系、相互促进的特殊矛盾性。毛泽东同志曾指出：“研究所有这些矛盾的特性……必须对它们实行具体的分析。离开具体的分析，就不能认识任何矛盾的特性”①。现在主要是在探讨统计方法的应用问题，才深刻地理解到毛泽东同志在1957年 3 月在中国共产党全国宣传工作会议上的讲话中所谈到的：“分析的方法就是辩证的方法。所谓分析，就是分析事物的矛盾。”是以，作为科学研究方法指导思想的具体分析，主要是分析客观存在着的某一具体的特殊矛盾运动，其次才是具体的因果分析和定量分析（包括大量观察下的统计分析）。就人口出生性比例这一科学研究问题的具体分析为例，可以说：生育性比例的规律性乃是精子细胞的XY型染色体与卵子细胞的XX型染色体结合成为一对具体而特殊的矛盾的运动发展。

毛泽东同志在《矛盾论》中就指出过“科学对象”的问题。他说：“科学研究的区分，就是根据科学对象所具有的特殊的矛盾性。因此，对于某一现象的领域所特有的某一种矛盾的研究，就构成某一门科学的对象。”②根据这一辩证唯物主义的基本观点，应该先研究一下统计科学研究对象所具有的特殊矛盾性。如前所论，统计学应该分为理论统计学和应用统计学，理论统计学（数理统计学）具有研究对象，而应用统计学（数理统计方法）则只有应用对象。理论统计学又包括两部分：概率论与数理统计；概率论是统计科学本身的理论基础，而数理统计则为测定必

① 《毛泽东选集》第 1 卷，第292页。

② 同上，第284页。

然性数量特征的数学方法。是以，统计科学研究的矛盾性，从理论上来说，就是概率论研究的特殊的矛盾性；而理论统计学的研究对象则为客观大量现象中个别的偶然性与集体的必然性之间对立统一抽象的数量规律性。应用统计学作为一种科学研究的必要工具，其应用对象则为各个科学领域普遍存在的集体大量现象的数量差异性。而这种大量现象的数量差异性又是各种科学领域具体对象中必然性与偶然性对立统一规律的具体表现。所以，在理论上，统计分析主要是分析必然性与偶然性这一对特殊的矛盾性。从各种科学领域的科学研究来看，还应注意区别：统计理论研究的矛盾性，和应用统计学于观察大量现象数量差异性所反映的矛盾性，和科学研究所探索的具体规律中的矛盾性。后二者的区别与关系，从人口出生性比例规律性这一具体科学研究的实例，体现得很清楚。现在需要探讨的是：应用统计学在各个特殊科学领域的具体科学研究中探索某一具体规律性中所起的特殊作用问题。应当指出：统计研究方法（计量描述、分析描述、分析推论）只是科学研究在大量观察下，按照必然性和偶然性对立统一的规律，分析必然性数量特征、关系、变化的一套统计分析工具。它既不等于进行定量分析的所有数学方法，更不能代替科学研究的方法论。科学研究的方法论应当从实际出发，了解并分析具体情况，探索特殊的矛盾性；然后在对具体情况具体分析特殊矛盾运动的基础上，联系有关研究对象的具体的理论分析（如人口出生性比例的生物学分析），进行具体的统计分析，才能探索到质与量相结合的具体规律性（如人口出生性比例的规律性）。

综合以上，统计方法（数理统计方法）只是科学研究中，分析客观大量现象中偶然性、测定必然性数量特征、数量关系、数量变化的一套分析工具。科学研究的方法必须在马克思主义“具体地分析具体的情况”这一方法论原理的指导下，在应用矛盾论的基础上，把统计分析及有关的数学分析与研究对象的具体科学



的理论（如经济学、人口学、生物学等）分析，按照质与量的辩证关系结合起来。可以说：科学研究工作本身就是在辩证唯物主义指导下的一种主观认识客观规律性的活动。所以，首先，科学研究从实际出发所观察或调查到的，只是反映种种客观具体现象的大量材料，这就需要“去粗取精、去伪存真”①，选择到能够如实反映客观具体情况的具体材料。其次，科学研究在分析和研究这些材料的过程中，就需要“由此及彼”②地寻找联系，分析矛盾的特性和主要矛盾着的两方面的具体情况，探索特殊矛盾运动的具体过程。同时，科学研究在具体的分析、研究过程中，不论在搜集事实材料方面，还是在对具体情况进行具体分析（包括特殊的矛盾分析、定量的统计分析和具体的数学分析、定性的具体科学理论分析）方面，都需要“由表及里”③，反复研究，层层深入，并且经过多次反复的科学实践，才能探索到客观存在着的某一具体规律性。

---

①②③ 《实践论》，《毛泽东选集》第1卷，第268页。